INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/29824

Box I Observation where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-23 ALL PARTLY

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-23 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 1, 2, 23, 24.

2. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 3, 4.

3. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 5, 6.

4. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 7, 8.

5. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 9, 10.

6. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 11, 12.

7. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 13, 14.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 15, 16.

9. Claims: 1-23 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 17, 18, 25, 26.

10. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 19, 20.

11. Claims: 1-23 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 21, 22, 27, 28.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/29824

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/54 C12N9/12 C07K14/415 C12N5/10 C12Q1/68
C12Q1/48

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SANO ET AL.: "Light and nutritional regulation of transcripts encoding a wheat protein kinase homolog is mediated by cytokinins" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 91, March 1994 (1994-03), pages 2582-2586, XP002137633 cited in the application abstract; figure 1 --- -/--	1,2,10, 22,23

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 May 2000	16. 8. 00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	CEDER O.

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

	Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
8	A	HOTTA H ET AL: "Molecular analysis of a novel protein kinase in maturing rice seed" GENE, vol. 213, no. 1-2, 15 June 1998 (1998-06-15), pages 47-54, XP004125002 ISSN: 0378-1119 abstract	5,14,15, 18

1	A	TSAI ET AL.: "SNF1-related protein kinase" EMBL SEQUENCE DATABASE, 1 November 1996 (1996-11-01), XP002137634 HEIDELBERG DE Ac Q40740 the whole document	1

1	A	TSAI ET AL.: "Oryza sativa SNF1-related protein kinase (RSk1) mRNA, complete cds" EMBL SEQUENCE DATABASE, 3 August 1996 (1996-08-03), XP002137635 HEIDELBERG DE Ac U55768 the whole document	1

1	A	ALDERSON ET AL.: "Complementation of SNF1, a mutation affecting global regulation of carbon metabolim in yeast, by a plant protein kinase cDNA" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 88, October 1991 (1991-10), pages 8602-8605, XP002137636 cited in the application abstract; figure 1	1

1	A	ANNEN ET AL.: "Characterization of 14 different putative protein kinase cdNA clones of the C4 plant Sorghum bicolor" EMBL SEQUENCE DATABASE, 18 August 1998 (1998-08-18), XP002137637 HEIDELBERG DE Ac AA738543 the whole document & MOL. GEN. GENET., vol. 259, no. 1, 1998, pages 115-122,	1

		-/--	
1			

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
1 A	MAN ET AL.: "Potato SNF1-related protein kinase: molecular cloning , expression, analysis and peptide kinase activity measurements" EMBL SEQUENCE DATABASE, 1 November 1996 (1996-11-01), XP002137638 HEIDELBERG DE Ac Q41485 the whole document & PLANT MOL BIOL, vol. 34, 1997, pages 31-43, ---	1
1 A	GUMPEL: "C. sativus mRNA dor SNF1-related protein kinase" EMBL SEQUENCE DATABASE, 17 December 1996 (1996-12-17), XP002137639 HEIDELBERG DE Ac Y10036 the whole document -----	1

1

INTERNATIONAL SEARCH REPORT

Inte...ational application No.
PCT/US 99/29824

Box I Observations where certain claims were found unsearchable (Continuation of item 1 f first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-23 ALL PARTLY

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-23 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 1, 2, 23, 24.

2. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 3, 4.

3. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 5, 6.

4. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 7, 8.

5. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 9, 10.

6. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 11, 12.

7. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 13, 14.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 15, 16.

9. Claims: 1-23 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 17, 18, 25, 26.

10. Claims: 1-21 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 19, 20.

11. Claims: 1-23 all partly

Isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression and their selection and uses where the sequences are seq id nos 21, 22, 27, 28.

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

SCHAEFFER, Andrew L.
DU PONT PHARMACEUTICALS COMPANY
Legal Patent Records Center
1007 Market Street
Wilmington, Delaware 19898
ETATS-UNIS D'AMERIQUE

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT
(PCT Rule 71.1)

Date of mailing *(day/month/year)*	23.02.2001
Applicant's or agent's file reference: BB1315	**IMPORTANT NOTIFICATION**

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/29824	15/12/1999	16/12/1998

Applicant
E.I. DU PONT DE NEMOURS AND COMPANY et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/	Authorized officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 pmu d Fax: +49 89 2399 - 4465	Hingel, W Tel.+49 89 2399-8717

Form PCT/IPEA/416 (July 1992)

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
BB1315		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/29824	15/12/1999	16/12/1998
International Patent Classification (IPC) or national classification and IPC C12N15/54		
Applicant E.I. DU PONT DE NEMOURS AND COMPANY et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 9 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
10/07/2000	23.02.2001
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Paresce, D Telephone No. +49 89 2399 8995

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/US99/29824

I. Basis of the report

1. This report has been drawn on the basis of *(substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments (Rules 70.16 and 70.17).)*:
Description, pages:

1-26 as originally filed

Claims, No.:

1-23 as originally filed

Sequence listing part of the description, pages:

1-28, as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☒ contained in the international application in written form.
- ☒ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

   (Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

   ☐ the entire international application.

   ☒ claims Nos. 1-23 partially.

because:

   ☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):

   ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):

   ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

   ☒ no international search report has been established for the said claims Nos. 1-23 partially.

2. A meaningful international preliminary examination report cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

   ☐ the written form has not been furnished or does not comply with the standard.

   ☐ the computer readable form has not been furnished or does not comply with the standard.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-23 partially
	No:	Claims	
Inventive step (IS)	Yes:	Claims	

	No:	Claims	1-23 partially
Industrial applicability (IA)	Yes:	Claims	1-23 partially
	No:	Claims	

2. Citations and explanations
see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:
see separate sheet

R It m I

Basis of the report

This International Preliminary Examination Report will be based on the subject-matter of claims 1-23 partially (see below).

Re Item III

Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The Examining Division agrees with the objection put forward by the Search Division as to lack of unity (Rule 13.1 PCT). As the Applicant has not had a search report drawn up on the other invention (Rule 68.5 PCT), the application will be prosecuted on the basis of the invention in respect of which a search has already been carried out, in other words, the invention first mentioned in the claims. This includes the subject-matter of claims 1-23 partially. These claims are directed to isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression, wherein the sequences are SEQ ID NOs: 1, 2, 23 and 24. The claims are further directed to the methods of selection and uses of said sequences (see International Search Report).

The Applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The Applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report of during any Chapter II procedure. The applicant should therefore, limit the application to the invention searched and excise those parts of the application relating to the other inventions.

R It m V

R as n d stat m nt under Articl 35(2) with r gard to nov lty, inv ntiv st p or

industrial applicability; citations and xplanations supporting such statement

The documents mentioned in this communication are numbered as in the search report, i.e. D1 corresponds to the first document of the search report.

The invention of the present application is the provision of nucleic acid sequences encoding all or portions of a corn sucrose non-fermenting protein (SNF1). At the priority date of the present application, SNF1 was known to play a central role in carbon catabolite repression in yeast. It was identified by screening for genes involved in the regulation of invertase gene expression and is required for the expression of several glucose-repressible genes in response to glucose deprivation (see D5, page 8604). In yeast, the SNF1 gene encodes a 72 kDa serine and threonine protein kinase whose activity is down regulated by the availability of glucose. In addition, several plant SNF1-like genes had been identified in the prior art. D1 discloses a wheat SNF1 protein kinase and its nucleotide and deduced amino acid sequence (D1, figure 1). The amino acid sequence of the protein kinase was compared to members of the SNF1 kinase subfamily. D2 and D3 disclose a rice SNF1-related protein kinase nucleotide and amino acid sequence. The polypeptide sequence disclosed in D3 shows 87.1% identity in a 510 amino acid overlap to SEQ ID NO: 4 of the present application. D6, D7 and D8 disclose SNF1 protein kinase polynucleotide and polypeptide sequences from (respectively) sorghum, potato, and cucumber. D5 describes a rye SNF1 protein kinase polynucleotide and polypeptide sequence (see figure 1). The polypeptide sequence disclosed in D5 shows 77.9% identity in a 510 amino acid overlap to SEQ ID NO: 4 of the present application. Expression of the rye SNF1 cDNA in yeast snf1 mutants restored SNF1 function, therefore, it was suggested that yeast and rye may have similar mechanisms for carbon catabolite repression (see abstract).

The subject-matter of claims 1-23 has not been made available to the public by any of the available prior art documents and can therefore be regarded as novel.

In view of the prior art, the problem to be solved by the present invention may be regarded as the provision of a further sequence encoding a plant SNF1-like protein. The solution, however, proposed in claims 1-23 of the present application

cannot be considered as involving an inventive step (Article 33(3) PCT) for the following reasons.

The polypeptide encoded by SEQ ID NO: 1 or 23 is characterized in structural terms but not in functional terms. Claim 1 is directed to a polypeptide characterized only by an amino acid sequence. Claim 2 is directed to a DNA molecule characterized only by a nucleic acid sequence. The DNA or polypeptide is thus structurally characterized but the claims include no other technical effect or function for said DNA.

According to page 2 of the description of the present application, SEQ ID NOs 1, 2, 23 and 24 are nucleic acid or amino acid sequences encoding a corn SNF1 protein. This was established by computer-assisted sequence alignments in which said sequences were compared with sequences of proteins of the SNF1 family. In fact, the sequences claimed in the present application were identified by standard procedures well-known to a person skilled in the art. cDNA libraries were prepared from various plant tissues and cDNA clones encoding SNF1-like proteins were identified by conducting BLAST. The cDNA sequences were compared to all publicly available DNA sequences and the nucleic acid fragments encoding probable SNF1 proteins were identified. From the sequence alignments with proteins of the SNF1 family it is concluded that, "these sequences represent the first corn, soybean and wheat sequences and a new rice sequence encoding SNF1 proteins" (present application, p. 21). At this stage, however, the protein is still only structurally defined. The comparison studies do not allow any concrete conclusions to be drawn concerning the actual function of the corn "SNF1", but only speculations that need to be confirmed experimentally.

If a nucleic acid or protein is only defined as being related or as having homology to another one, it has no definite function. Even though the corn SNF1 is defined in the present application as belonging to the SNF1 family, different members of any family can have very different functions, for example see pages 1-2 of the description. Consequently, such a definition cannot be considered to contain any technical information. Without any further indication of which of the many functions of the SNF family are intended for the claimed protein, the skilled person would be left guessing as to what the claimed protein should be used for. Examples 4-6 of

the present application merely describe methods of expressing the claimed cDNA in various host cells. These examples provide no further information on the function or technical features of the claimed protein.

If an invention should provide a solution to a problem with reference to the background art (Article 33(3) PCT), the invention of the present application is insufficiently disclosed and unclear (Article 6 PCT) since it is left to the reader to perform the invention and to determine what problem if any, the isolated nucleotide or polypeptide sequences actually solve.

Furthermore, due to the fact that the claimed sequences are not associated with any known technical effect, the only problem which might be recognized is the mere provision of further DNA sequences **as such**, regardless of their possible useful properties. In this case all known DNA sequences are equally suitable candidates for solving the above "technical problem" and would, therefore, all equally be suggested to the skilled person. The arbitrary selection from an infinite number of equally obvious possible solutions cannot involve an inventive step because, in order to be patentable, the selection must not be arbitrary but must be justified by the technical purpose, e.g. by a hitherto unknown or unexpected technical effect which is caused by those **structural features** distinguishing the claimed compounds from the numerous other ones.

Furthermore, procedures for screening and cloning using EST sequences were routine at the priority date of the present application (16.12.98). Therefore, in general, the skilled person could be expected to perform the cloning of a gene with a reasonable expectation of success and without undue burden. Furthermore, procedures for performing sequence alignments and production of expression systems are all standard procedures well known to a person skilled in the art. Therefore, the IPEA is of the opinion that the polynucleotide or polypeptide sequences in question cannot be considered to fulfill the requirements of Article 6 PCT with regard to clarity as well as the requirements of Article 33(3) PCT with respect to inventive step.

Furthermore, in so far as the present specification is completely silent with regard to the function of the protein encoded by SEQ ID NOs:1, 2, 23, 24, the possible

fields of application for the claimed protein and polynucleotides suggested in the specification are only speculative. Thus it is questionable whether the claimed subject-matter actually meets the requirements set forth in Article 33(4) PCT.

VIII. Certain observations on the international application

1) Clarity: Article 6 PCT

Article 6 PCT requires amongst other things that the claims, which define the matter for which protection is sought (i.e. the object of invention) be clear. This has to be interpreted as meaning not only that a claim from a technical point of view must be comprehensible, but also that it must define clearly the object of the invention, that is to say, it must indicate all the essential features thereof. In the present case, claim 1 refers to a polypeptide comprising a given sequence. An amino acid sequence, however, does not provide sufficient information to clearly characterize the claimed polypeptide. The polypeptide in question should be additionally characterized by name, source and function. Claim 2 at present covers DNA sequences which encode a polypeptide without any biological function. These claims should also include the function of said polypeptide. However, the attention of the applicant is directed to the fact that such a functional limitation must have a basis in the disclosure of the present application.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
BB1315		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/29824	15/12/1999	16/12/1998
International Patent Classification (IPC) or national classification and IPC C12N15/54		
Applicant E.I. DU PONT DE NEMOURS AND COMPANY et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 9 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
10/07/2000	23.02.2001
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Paresce, D Telephone No. +49 89 2399 8995

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/US99/29824

I. Basis f th r p rt

1. This report has been drawn on the basis of *(substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments (Rules 70.16 and 70.17).)*:
Description, pages:

1-26 as originally filed

Claims, No.:

1-23 as originally filed

Sequence listing part of the description, pages:

1-28, as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☒ contained in the international application in written form.
- ☒ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the clairns, Nos.:
- ☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

   (Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

   ☐ the entire international application.

   ☒ claims Nos. 1-23 partially.

because:

   ☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):

   ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):

   ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

   ☒ no international search report has been established for the said claims Nos. 1-23 partially.

2. A meaningful international preliminary examination report cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

   ☐ the written form has not been furnished or does not comply with the standard.

   ☐ the computer readable form has not been furnished or does not comply with the standard.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-23 partially
	No:	Claims	
Inventive step (IS)	Yes:	Claims	

	No:	Claims	1-23 partially
Industrial applicability (IA)	Yes:	Claims	1-23 partially
	No:	Claims	

2. Citations and explanations
see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:
see separate sheet

Re Item I

Basis of the report

This International Preliminary Examination Report will be based on the subject-matter of claims 1-23 partially (see below).

Re Item III

Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The Examining Division agrees with the objection put forward by the Search Division as to lack of unity (Rule 13.1 PCT). As the Applicant has not had a search report drawn up on the other invention (Rule 68.5 PCT), the application will be prosecuted on the basis of the invention in respect of which a search has already been carried out, in other words, the invention first mentioned in the claims. This includes the subject-matter of claims 1-23 partially. These claims are directed to isolated polynucleotide and polypeptide sequences encoding a plant protein (SNF1 protein kinase) involved in carbon catabolite repression, wherein the sequences are SEQ ID NOs: 1, 2, 23 and 24. The claims are further directed to the methods of selection and uses of said sequences (see International Search Report).

The Applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The Applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report of during any Chapter II procedure. The applicant should therefore, limit the application to the invention searched and excise those parts of the application relating to the other inventions.

Re Item V

Reasoned statem nt under Article 35(2) with regard to novelty, inventiv st p or

industrial applicability; citations and xplanations supporting such stat ment

The documents mentioned in this communication are numbered as in the search report, i.e. D1 corresponds to the first document of the search report.

The invention of the present application is the provision of nucleic acid sequences encoding all or portions of a corn sucrose non-fermenting protein (SNF1). At the priority date of the present application, SNF1 was known to play a central role in carbon catabolite repression in yeast. It was identified by screening for genes involved in the regulation of invertase gene expression and is required for the expression of several glucose-repressible genes in response to glucose deprivation (see D5, page 8604). In yeast, the SNF1 gene encodes a 72 kDa serine and threonine protein kinase whose activity is down regulated by the availability of glucose. In addition, several plant SNF1-like genes had been identified in the prior art. D1 discloses a wheat SNF1 protein kinase and its nucleotide and deduced amino acid sequence (D1, figure 1). The amino acid sequence of the protein kinase was compared to members of the SNF1 kinase subfamily. D2 and D3 disclose a rice SNF1-related protein kinase nucleotide and amino acid sequence. The polypeptide sequence disclosed in D3 shows 87.1% identity in a 510 amino acid overlap to SEQ ID NO: 4 of the present application. D6, D7 and D8 disclose SNF1 protein kinase polynucleotide and polypeptide sequences from (respectively) sorghum, potato, and cucumber. D5 describes a rye SNF1 protein kinase polynucleotide and polypeptide sequence (see figure 1). The polypeptide sequence disclosed in D5 shows 77.9% identity in a 510 amino acid overlap to SEQ ID NO: 4 of the present application. Expression of the rye SNF1 cDNA in yeast snf1 mutants restored SNF1 function, therefore, it was suggested that yeast and rye may have similar mechanisms for carbon catabolite repression (see abstract).

The subject-matter of claims 1-23 has not been made available to the public by any of the available prior art documents and can therefore be regarded as novel.

In view of the prior art, the problem to be solved by the present invention may be regarded as the provision of a further sequence encoding a plant SNF1-like protein. The solution, however, proposed in claims 1-23 of the present application

cannot be considered as involving an inventive step (Article 33(3) PCT) for the following reasons.

The polypeptide encoded by SEQ ID NO: 1 or 23 is characterized in structural terms but not in functional terms. Claim 1 is directed to a polypeptide characterized only by an amino acid sequence. Claim 2 is directed to a DNA molecule characterized only by a nucleic acid sequence. The DNA or polypeptide is thus structurally characterized but the claims include no other technical effect or function for said DNA.

According to page 2 of the description of the present application, SEQ ID NOs 1, 2, 23 and 24 are nucleic acid or amino acid sequences encoding a corn SNF1 protein. This was established by computer-assisted sequence alignments in which said sequences were compared with sequences of proteins of the SNF1 family. In fact, the sequences claimed in the present application were identified by standard procedures well-known to a person skilled in the art. cDNA libraries were prepared from various plant tissues and cDNA clones encoding SNF1-like proteins were identified by conducting BLAST. The cDNA sequences were compared to all publicly available DNA sequences and the nucleic acid fragments encoding probable SNF1 proteins were identified. From the sequence alignments with proteins of the SNF1 family it is concluded that, "these sequences represent the first corn, soybean and wheat sequences and a new rice sequence encoding SNF1 proteins" (present application, p. 21). At this stage, however, the protein is still only structurally defined. The comparison studies do not allow any concrete conclusions to be drawn concerning the actual function of the corn "SNF1", but only speculations that need to be confirmed experimentally.

If a nucleic acid or protein is only defined as being related or as having homology to another one, it has no definite function. Even though the corn SNF1 is defined in the present application as belonging to the SNF1 family, different members of any family can have very different functions, for example see pages 1-2 of the description. Consequently, such a definition cannot be considered to contain any technical information. Without any further indication of which of the many functions of the SNF family are intended for the claimed protein, the skilled person would be left guessing as to what the claimed protein should be used for. Examples 4-6 of

the present application merely describe methods of expressing the claimed cDNA in various host cells. These examples provide no further information on the function or technical features of the claimed protein.

If an invention should provide a solution to a problem with reference to the background art (Article 33(3) PCT), the invention of the present application is insufficiently disclosed and unclear (Article 6 PCT) since it is left to the reader to perform the invention and to determine what problem if any, the isolated nucleotide or polypeptide sequences actually solve.

Furthermore, due to the fact that the claimed sequences are not associated with any known technical effect, the only problem which might be recognized is the mere provision of further DNA sequences **as such**, regardless of their possible useful properties. In this case all known DNA sequences are equally suitable candidates for solving the above "technical problem" and would, therefore, all equally be suggested to the skilled person. The arbitrary selection from an infinite number of equally obvious possible solutions cannot involve an inventive step because, in order to be patentable, the selection must not be arbitrary but must be justified by the technical purpose, e.g. by a hitherto unknown or unexpected technical effect which is caused by those **structural features** distinguishing the claimed compounds from the numerous other ones.

Furthermore, procedures for screening and cloning using EST sequences were routine at the priority date of the present application (16.12.98). Therefore, in general, the skilled person could be expected to perform the cloning of a gene with a reasonable expectation of success and without undue burden. Furthermore, procedures for performing sequence alignments and production of expression systems are all standard procedures well known to a person skilled in the art. Therefore, the IPEA is of the opinion that the polynucleotide or polypeptide sequences in question cannot be considered to fulfill the requirements of Article 6 PCT with regard to clarity as well as the requirements of Article 33(3) PCT with respect to inventive step.

Furthermore, in so far as the present specification is completely silent with regard to the function of the protein encoded by SEQ ID NOs:1, 2, 23, 24, the possible

fields of application for the claimed protein and polynucleotides suggested in the specification are only speculative. Thus it is questionable whether the claimed subject-matter actually meets the requirements set forth in Article 33(4) PCT.

VIII. Certain observations on the international application

1) Clarity: Article 6 PCT

Article 6 PCT requires amongst other things that the claims, which define the matter for which protection is sought (i.e. the object of invention) be clear. This has to be interpreted as meaning not only that a claim from a technical point of view must be comprehensible, but also that it must define clearly the object of the invention, that is to say, it must indicate all the essential features thereof. In the present case, claim 1 refers to a polypeptide comprising a given sequence. An amino acid sequence, however, does not provide sufficient information to clearly characterize the claimed polypeptide. The polypeptide in question should be additionally characterized by name, source and function. Claim 2 at present covers DNA sequences which encode a polypeptide without any biological function. These claims should also include the function of said polypeptide. However, the attention of the applicant is directed to the fact that such a functional limitation must have a basis in the disclosure of the present application.

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 11 August 2000 (11.08.00)	
International application No. PCT/US99/29824	**Applicant's or agent's file reference** BB1315 PCT
International filing date (day/month/year) 15 December 1999 (15.12.99)	**Priority date** (day/month/year) 16 December 1998 (16.12.98)
Applicant ALLEN, Stephen, M. et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

10 July 2000 (10.07.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer F. Baechler
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/54, 9/12, C07K 14/415, C12N 5/10, C12Q 1/68, 1/48

A2

(11) International Publication Number: **WO 00/36115**

(43) International Publication Date: 22 June 2000 (22.06.00)

(21) International Application Number: PCT/US99/29824

(22) International Filing Date: 15 December 1999 (15.12.99)

(30) Priority Data:
60/112,563 16 December 1998 (16.12.98) US

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ALLEN, Stephen, M. [US/US]; 2225 Rosewood Drive, Wilmington, DE 19810 (US). HEPPARD, Elmer, P. [US/US]; 613 Governor Circle, Wilmington, DE 19809 (US). SAKAI, Hajime [DE/US]; 105 Banbury Drive, Wilmington, DE 19803 (US). WENG, Zude [CN/US]; 909C Cloister Road, Wilmington, DE 19809 (US). HELENTJARIS, Timothy, G. [US/US]; 2960 N.W. 73rd Lane, Ankeny, IA 50021 (US). MACOOL, Daniel, J. [US/US]; 2525 Meredith Street, Philadelphia, PA 19130 (US). MIAO, Guo-Hua [US/US]; 202 Cherry Blossom Place, Hockessin, DE 19707 (US).

(74) Agent: FEULNER, Gregory, J.; E.I. du Pont de Nemours and Company, Legal Patent Record Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) **Title:** PLANT CARBON CATABOLITE REPRESSION PROTEINS

(57) **Abstract**

This invention relates to an isolated nucleic acid fragment encoding a carbon catabolite repression polypeptide. The invention also relates to the construction of a chimeric gene encoding all or a portion of the carbon catabolite repression polypeptide, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the carbon catabolite repression polypeptide in a transformed host cell.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE

PLANT CARBON CATABOLITE REPRESSION PROTEINS

This application claims the benefit of U.S. Provisional Application No. 60/112,563, filed December 16, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding proteins involved in carbon catabolite repression in plants and seeds.

BACKGROUND OF THE INVENTION

Improvement of crop plants for a variety of traits, including disease and pest resistance, and grain quality improvements such as oil, starch or protein composition are major goals of plant breeders. Conventional plant breeding will continue to provide improved crops, however metabolic engineering through biotechnology may provide the greatest potential for crop improvement.

The major storage compounds found in most plants are carbohydrates, lipids and proteins and the relative abundance and proportions of these compounds vary between plants and in plant storage and vegetative organs. Lipids and proteins are the major storage compounds in peanuts and soybeans, whereas carbohydrate is the predominant form in cereals. Specific mechanisms control the allocation of photoassimilates into the carbon and nitrogen metabolic pathways that produce these storage compounds, however, little is known about the regulation of these pathways. Many of the plant genes that are involved in photosynthesis and metabolic pathways appear to be regulated by catabolite repression mechanisms (Sheen, J. (1990) *Plant Cell 2*:2107-2138). Catabolite repression is a general mechanism utilized by prokaryotes and lower eukaryotes to regulate carbon and nitrogen metabolism. Plant homologs of several proteins involved in catabolite regulation of various metabolic pathways in bacterial and yeast suggest that similar regulatory mechanisms exist in plants (Alderson, A. et al., (1991) *PNAS 88*:8602-8605 and Sano, H. et al., (1994) *PNAS 91*:2582-2586).

In yeast, carbon sensing and glucose repression is regulated by a cAMP independent pathway. One component of the glucose repression pathway is the sucrose non-fermenting (SNF1) protein. SNF1 is a global regulatory protein that is required for the expression of glucose-repressible genes. The SNF1 gene encodes a 72 kDa serine and threonine protein kinase whose activity is down regulated by the availability of glucose (Celenza, J. L. et al., (1986) *Science 233*:1175-1180). The PHO85 gene of yeast encodes a glycogen synthase kinase that also appears to be involved in the glucose repression pathway possibly acting downstream of SNF1. Huang et al. report that cells with disrupted PHO85 genes accumulated large quantities of glycogen, had an active glycogen synthase and were impaired in glycogen synthase kinase activity. In SNF1 mutant cells glycogen synthase

activity was elevated suggesting that PHO85 and SNF1 act antagonistically to control glycogen accumulation (Huang, D. et al., (1996) *Mol Cell Biol 16(8)*:4357-4365).

Several plant SNF1-like genes have been identified which suggests that these plant genes may constitute a large family of protein kinases involved in carbon catabolite regulation. One rye SNF1-like cDNA homolog, cRKINI, has been shown to be capable of rescuing SNF1 mutations in yeast cells suggesting that yeast and rye may have similar mechanisms of regulating carbon metabolism (Miao, G. H. (1996) in *Plant Gene Research: Basic Knowledge and Application* ed. by D. P. S. Verma, Springer-Verlag, New York). Thus there is a great deal of interest in identifying the genes that encode proteins involved in catabolite repression in plants. These genes may be used in plants to control carbon and nitrogen partitioning pathways during plant growth and development. Accordingly, the availability of nucleic acid sequences encoding all or a portion of SNF1 and PHO85 regulatory proteins would facilitate studies to better understand the mechanism of catabolite repression in plants and could provide genetic tools to enhance or otherwise alter the accumulation of carbohydrates, lipids and proteins during plant growth and development.

SUMMARY OF THE INVENTION

The present invention relates to isolated polynucleotides comprising a nucleotide sequence encoding a polypeptide of at least 240 amino acids that has at least 90% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of corn SNF1 polypeptides of SEQ ID NOs:2, 4 and 6, a rice SNF1 polypeptide of SEQ ID NO:8, soybean SNF1 polypeptides of SEQ ID NOs:10, 12, 14, 16 and 18, wheat SNF1 polypeptides of SEQ ID NOs:20 and 22. The present invention also relates to an isolated polynucleotide comprising the complement of the nucleotide sequences described above.

The present invention also relates to isolated polynucleotides comprising a nucleotide sequence encoding a polypeptide of at least 77 amino acids that has at least 85% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of a corn SNF1 polypeptide of SEQ ID NO:24, a soybean SNF1 polypeptide of SEQ ID NO:26, and a wheat SNF1 polypeptide of SEQ ID NO:28. The present invention also relates to an isolated polynucleotide comprising the complement of the nucleotide sequences described above.

It is preferred that the isolated polynucleotides of the claimed invention consists of a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27 that codes for the polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26 and 28. The present invention also relates to an isolated polynucleotide comprising a nucleotide sequences of at least one of 60 (preferably at least one of 40, most preferably at least one of 30) contiguous

nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27 and the complement of such nucleotide sequences.

The present invention relates to a chimeric gene comprising an isolated polynucleotide of the present invention operably linked to suitable regulatory sequences.

The present invention relates to an isolated host cell comprising a chimeric gene of the present invention or an isolated polynucleotide of the present invention. The host cell may be eukaryotic, such as a yeast or a plant cell, or prokaryotic, such as a bacterial cell. The present invention also relates to a virus, preferably a baculovirus, comprising an isolated polynucleotide of the present invention or a chimeric gene of the present invention.

The present invention relates to a process for producing an isolated host cell comprising a chimeric gene of the present invention or an isolated polynucleotide of the present invention, the process comprising either transforming or transfecting an isolated compatible host cell with a chimeric gene or isolated polynucleotide of the present invention.

The present invention relates to a SNF1 polypeptide of at least 240 amino acids comprising at least 90% homology based on the Clustal method of alignment compared to a polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22.

The present invention relates to a SNF1 polypeptide of at least 77 amino acids comprising at least 85% homology based on the Clustal method of alignment compared to a polypeptide selected from the group consisting of SEQ ID NOs:24, 26 and 28.

The present invention relates to a method of selecting an isolated polynucleotide that affects the level of expression of a SNF1 polypeptide in a host cell, preferably a plant cell, the method comprising the steps of: (a) constructing an isolated polynucleotide of the present invention or an isolated chimeric gene of the present invention; (b) introducing the isolated polynucleotide or the isolated chimeric gene into a host cell; (c) measuring the level a SNF1 polypeptide in the host cell containing the isolated polynucleotide; and (d) comparing the level of a SNF1 polypeptide in the host cell containing the isolated polynucleotide with the level of a SNF1 polypeptide in the host cell that does not contain the isolated polynucleotide.

The present invention relates to a method of obtaining a nucleic acid fragment encoding a substantial portion of a SNF1 polypeptide gene, preferably a plant SNF1 polypeptide gene, comprising the steps of: synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 60 (preferably at least one of 40, most preferably at least one of 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27 and the complement of such nucleotide sequences; and amplifying a nucleic acid fragment (preferably a cDNA inserted in a cloning vector) using the oligonucleotide primer.

The amplified nucleic acid fragment preferably will encode a portion of a SNF1 amino acid sequence.

The present invention also relates to a method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a SNF1 polypeptide comprising the steps of: probing a cDNA or genomic library with an isolated polynucleotide of the present invention; identifying a DNA clone that hybridizes with an isolated polynucleotide of the present invention; isolating the identified DNA clone; and sequencing the cDNA or genomic fragment that comprises the isolated DNA clone.

The present invention relates to a composition, such as a hybridization mixture, comprising an isolated polynucleotide of the present invention.

The present invention relates to an isolated polynucleotide of the present invention comprising at least one of 30 contiguous nucleotides derived from a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27.

The present invention relates to an expression cassette comprising an isolated polynucleotide of the present invention operably linked to a promoter.

The present invention relates to a method for positive selection of a transformed cell comprising: (a) transforming a host cell with the chimeric gene of the present invention or an expression cassette of the present invention; and (b) growing the transformed host cell, preferably plant cell, such as a monocot or a dicot, under conditions which allow expression of the SNF1 polynucleotide in an amount sufficient to complement a null mutant and/or alter expression of glucose repressible genes to provide a positive selection means.

BRIEF DESCRIPTION OF THE SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying Sequence Listing which form a part of this application.

Table 1 lists the polypeptides that are described herein, the designation of the cDNA clones that comprise the nucleic acid fragments encoding polypeptides representing all or a substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as used in the attached Sequence Listing. Table 1 also identifies the cDNA clones as individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), contigs assembled from two or more ESTs ("Contig"), contigs assembled from an FIS and one or more ESTs ("Contig*"), or sequences encoding the entire protein derived from an FIS, a contig, or an FIS and PCR ("CGS"). Nucleotide sequences, SEQ ID NOs:1, 17 and 21 and amino acid sequences SEQ ID NOs:2, 18 and 22 were determined by further sequence analysis of cDNA clones encoding the amino acid sequences set forth in SEQ ID NOs:24, 26 and 28. Nucleotide SEQ ID NOs:23, 25 and 27 and amino acid SEQ ID NOs:24, 26 and 28 were presented in a U.S. Provisional Application No. 60/112,563, filed December 16, 1998.

The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

TABLE 1

Proteins Involved In Carbon Catabolite Repression

Protein	Clone Designation	SEQ ID NO: (Nucleotide)	SEQ ID NO: (Amino Acid)
SNF1	cen3n.pk0044.b8 (CGS)	1	2
SNF1	p0016.ctsbf56rb (CGS)	3	4
SNF1	p0118.chsbh89r (CGS)	5	6
SNF1	Contig composed of: rda.pk0007.g3 rr1.pk0008.e12 rr1.pk0047.g12	7	8
SNF1	sdp4c.pk016.e10 (CGS)	9	10
SNF1	sdr1f.pk001.p7 (CGS)	11	12
SNF1	sgs4c.pk006.g6 (CGS)	13	14
SNF1	sgs4c.pk006.n21 (CGS)	15	16
SNF1	srr1c.pk001.i24 (GCS)	17	18
SNF1	Contig composed of: wdk2c.pk018.c16 wlm1.pk0027.a12	19	20
SNF1	wlm96.pk0007.e4 (CGS)	21	22
	cen3n.pk0044.b8 (EST)	23	24
	srr1c.pk001.i24 (EST)	25	26
	wlm96.pk0007.e4 (EST)	27	28

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res. 13*:3021-3030 (1985) and in the *Biochemical J. 219 (No. 2)*:345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a "polynucleotide" is a nucleotide sequence such as a nucleic acid fragment. A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA,

synthetic DNA, or mixtures thereof. An isolated polynucleotide of the present invention may include at least one of 60 contiguous nucleotides, preferably at least one of 40 contiguous nucleotides, most preferably one of at least 30 contiguous nucleotides derived from SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 or the complement of such sequences.

As used herein, "contig" refers to a nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of sequence homology. For example, the nucleotide sequences of two or more nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences (and thus their corresponding nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

Substantially similar nucleic acid fragments may be selected by screening nucleic acid fragments representing subfragments or modifications of the nucleic acid fragments of the instant invention, wherein one or more nucleotides are substituted, deleted and/or inserted, for their ability to affect the level of the polypeptide encoded by the unmodified nucleic acid fragment in a plant or plant cell. For example, a substantially similar nucleic acid fragment representing at least one of 30 contiguous nucleotides derived from the instant nucleic acid fragment can be constructed and introduced into a plant or plant cell. The level of the polypeptide encoded by the unmodified nucleic acid fragment present in a plant or plant cell exposed to the substantially similar nucleic fragment can then be compared to the level of the polypeptide in a plant or plant cell that is not exposed to the substantially similar nucleic acid fragment.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than

the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Consequently, an isolated polynucleotide comprising a nucleotide sequence of at least one of 60 (preferably at least one of 40, most preferably at least one of 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and the complement of such nucleotide sequences may be used in methods of selecting an isolated polynucleotide that affects the expression of a polypeptide (such as SNF1) in a host cell. A method of selecting an isolated polynucleotide that affects the level of expression of a polypeptide (such as SNF1) in a host cell (eukaryotic, such as plant or yeast, prokaryotic such as bacterial, or viral) may comprise the steps of: constructing an isolated polynucleotide of the present invention or an isolated chimeric gene of the present invention; introducing the isolated polynucleotide or the isolated chimeric gene into a host cell; measuring the level a polypeptide in the host cell containing the isolated polynucleotide; and comparing the level of a polypeptide in the host cell containing the isolated polynucleotide with the level of a polypeptide in a host cell that does not contain the isolated polynucleotide.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with

0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Suitable nucleic acid fragments (isolated polynucleotides of the present invention) encode polypeptides that are at least about 70% identical, preferably at least about 80% identical to the amino acid sequences reported herein. Preferred nucleic acid fragments encode amino acid sequences that are about 85% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are at least about 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are at least about 95% identical to the amino acid sequences reported herein. Suitable nucleic acid fragments not only have the above homologies but typically encode a polypeptide having at least about 50 amino acids, preferably at least about 100 amino acids, more preferably at least about 150 amino acids, still more preferably at least about 200 amino acids, and most preferably at least about 250 amino acids. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS. 5*:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization

of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. "Chemically synthesized", as related to nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature.

Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

"Coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

"Promoter" refers to a nucleotide sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a nucleotide sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants 15*:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Mol. Biotechnol. 3*:225-236).

The "3' non-coding sequences" refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell 1*:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into polypeptide by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA and so can be translated into a polypeptide by the cell. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

"Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol. 42*:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*) can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included (Raikhel (1992) *Plant Phys. 100*:1627-1632).

"Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol. 143*:277) and particle-accelerated or "gene gun" transformation technology (Klein et al. (1987) *Nature (London) 327*:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Maniatis").

Nucleic acid fragments encoding at least a portion of several proteins involved in carbon catabolite repression have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other SNF1 polypeptides, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA 85*:8998-9002) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA 86*:5673-5677; Loh et al. (1989) *Science 243*:217-220). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques 1*:165). Consequently, a polynucleotide comprising a nucleotide sequence of at least one of 60 (preferably one of at least 40, most preferably one of at least 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and the complement of such nucleotide sequences may be used in such methods to obtain a nucleic acid fragment encoding a substantial portion of an amino acid sequence of a polypeptide. The present invention relates to a method of obtaining a nucleic acid fragment encoding a substantial portion of a polypeptide of a gene (such as SNF1) preferably a substantial portion of a plant polypeptide of a gene, comprising the steps of: synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 60 (preferably at least one of 40,

most preferably at least one of 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and the complement of such nucleotide sequences; and amplifying a nucleic acid fragment (preferably a cDNA inserted in a cloning vector) using the oligonucleotide primer. The amplified nucleic acid fragment preferably will encode a portion of a polypeptide (SNF1).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol. 36*:1-34; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of catabolite repression in those cells.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then be constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al. (1985) *EMBO J. 4*:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics 218*:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant polypeptides to different cellular compartments, or to facilitate its secretion from the cell. It is thus envisioned that the chimeric gene described above may be further supplemented by directing the coding

sequence to encode the instant polypeptides with appropriate intracellular targeting sequences such as transit sequences (Keegstra (1989) *Cell 56*:247-253), signal sequences or sequences encoding endoplasmic reticulum localization (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol. 42*:21-53), or nuclear localization signals (Raikhel (1992) *Plant Phys.100*:1627-1632) with or without removing targeting sequences that are already present. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of use may be discovered in the future.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptides in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U.S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative regulation available with a transgenic approach may be advantageous from a breeding perspective. In addition, the ability to restrict the expression of specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppression technologies in order to reduce expression of particular genes. For example, the proper level of expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display the desired phenotype. Accordingly, the skilled artisan will develop methods for screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds, and is not an inherent part of the invention. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein

encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one which allows large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptides (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptides of the instant invention *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant polypeptides are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant polypeptides. This chimeric gene could then be introduced into appropriate microorganisms via transformation to provide high level expression of the encoded proteins involved in carbon catabolite repression. An example of a vector for high level expression of the instant polypeptides in a bacterial host is provided (Example 6).

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al. (1987) *Genomics 1*:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet. 32*:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter 4*:37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; *see* Hoheisel et al. In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask (1991) *Trends Genet. 7*:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan et al. (1995) *Genome Res. 5*:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med. 11*:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics 16*:325-332), allele-specific ligation (Landegren et al. (1988) *Science 241*:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res. 18*:3671), Radiation Hybrid Mapping (Walter et al. (1997) *Nat. Genet. 7*:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res. 17*:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA 86*:9402-9406; Koes et al. (1995) *Proc. Natl. Acad. Sci USA 92*:8149-8153; Bensen et al. (1995) *Plant Cell 7*:75-84). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptides. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With

either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptides can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptides disclosed herein.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various corn, rice, soybean and wheat tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from Corn, Rice, Soybean and Wheat

Library	Tissue	Clone
cen3n	Corn Endosperm 20 Days After Pollination*	cen3n.pk0044.b8
p0016	Corn Tassel Shoots, Pooled, 0.1-1.4 cm	p0016.ctsbf56rb
p0118	Corn Stem Tissue Pooled From the 4-5 Internodes Subtending The Tassel At Stages V8-V12**, Night harvested*	p0118.chsbh89r
rda	Rice 15 day old leaf*	rda.pk0007.g3
rr1	Rice Root of Two Week Old Developing Seedling	rr1.pk0008.e12 rr1.pk0047.g12
sdp4c	Soybean Developing Pods (10-12 mm)	sdp4c.pk016.e10
sdr1f		sdr1f.pk001.p7
sgs4c	Soybean Seeds 2 Days After Germination	sgs4c.pk006.g6 sgs4c.pk006.n21
srr1c	Soybean 8-Day-Old Root	srr1c.pk001.i24
wdk2c	Wheat Developing Kernel, 7 Days After Anthesis	wdk2c.pk018.c16

Library	Tissue	Clone
wlm1	Wheat Seedlings 1 Hour After Inoculation With *Erysiphe graminis f. sp tritici*	wlm1.pk0027.a12
wlm96	Wheat Seedlings 96 Hours After Inoculation With *Erysiphe graminis f. sp tritici*	wlm96.pk0007.e4

*These libraries were normalized essentially as described in U.S. Patent No. 5,482,845, incorporated herein by reference.

**Corn developmental stages are explained in the publication "How a corn plant develops" from the Iowa State University Coop. Ext. Service Special Report No. 48 reprinted June 1993.

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science 252*:1651-1656). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

EXAMPLE 2

Identification of cDNA Clones

cDNA clones encoding proteins involved in carbon catabolite repression were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information

(NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nat. Genet. 3*:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3

Characterization of cDNA Clones Encoding Proteins Involved in Carbon Catabolite Repression

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to SNF1 from *Arabidopsis thaliana* (NCBI Identifier No. gi 5051782), *Arabidopsis thaliana* (NCBI Identifier No. gi 4895200), *Arabidopsis thaliana* (NCBI Identifier No. gi 2980770), *Arabidopsis thaliana* (NCBI Identifier No. gi 3885328), *Cucumis sativus* (NCBI Identifier No. gi 1743009), *Glycine max* (NCBI Identifier No. gi 4567091), *Oryza sativa* (NCBI Identifier No. gi 4107001), *Oryza sativa* (NCBI Identifier No. gi 4107009). Shown in Table 3 are the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), contigs assembled from two or more ESTs ("Contig"), contigs assembled from an FIS and one or more ESTs ("Contig*"), or sequences encoding the entire protein derived from an FIS, a contig, or an FIS and PCR ("CGS"):

TABLE 3

BLAST Results for Sequences Encoding Polypeptides Homologous to *Arabidopsis thaliana*, *Cucumis sativus*, *Glycine max*, and *Oryza sativa* SNF1 Proteins

Clone	Status	BLAST pLog Score
cen3n.pk0044.b8	FIS	>254.00 (gi 5051782)
p0016.ctsbf56rb	CGS	>254.00 (gi 4107001)
p0118.chsbh89r	CGS	>254.00 (gi 4107009)
Contig composed of: rda.pk0007.g3 rr1.pk0008.e12 rr1.pk0047.g12	Contig	58.30 (gi 4895200)
sdp4c.pk016.e10	CGS	>254.00 (gi 2980770)
sdr1f.pk001.p7	CGS	>254.00 (gi 1743009)
sgs4c.pk006.g6	CGS	>254.00 (gi 2980770)
sgs4c.pk006.n21	CGS	>254.00 (gi 4567091)
srr1c.pk001.i24	FIS	>254.00 (gi 38853228)

Clone	Status	BLAST pLog Score
Contig composed of: wdk2c.pk018.c16 wlm1.pk0027.a12	Contig	158.00 (gi 4107001)
wlm96.pk0007.e4	FIS	>254.00 (4107009)

The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4. 6, 8, 10, 12, 14, 16, 18, 20 and 22 and the *Arabidopsis thaliana*, *Cucumis sativus*, *Glycine max*, and *Oryza sativa* sequences.

TABLE 4

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to *Arabidopsis thaliana*, *Cucumis sativus*, *Glycine max*, and *Oryza sativa* SNF1 Proteins

SEQ ID NO.	Percent Identity
2	73% (gi 5051782)
4	89% (gi 4107001)
6	88% (gi 4107009)
8	61% (gi 4895200)
10	74% (gi 2980770)
12	78% (gi 1743009)
14	72% (gi 2980770)
16	98% (gi 4567091)
18	82% (gi 3885328)
20	93% (gi 4107001)
22	81% (gi 4107009)

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS. 5*:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode a substantial portion of a SNF1 protein. These sequences represent the first corn, soybean and wheat sequences and a new rice sequence encoding SNF1 proteins.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding the instant polypeptides in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb SalI-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-SalI fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptides, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236)

which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* *8*:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem. 261*:9228-9238) can be used for expression of the instant polypeptides in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embryos may then be transformed with the expression vector comprising sequences encoding the instant polypeptides. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) *327*:70-73, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature 313*:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al.(1983) *Gene 25*:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression

cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptides and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 μL of a 60 mg/mL 1 μm gold particle suspension is added (in order): 5 μL DNA (1 μg/μL), 20 μl spermidine (0.1 M), and 50 μL $CaCl_2$ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μL 70% ethanol and resuspended in 40 μL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μL of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant polypeptides can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene 56*:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 μg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 μL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 μg/mL ampicillin. Transformants containing the gene encoding the instant polypeptides are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol. 189*:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One μg of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

Various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

The disclosure of each reference set forth above is incorporated herein by reference in its entirety.

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising a first nucleotide sequence encoding a polypeptide of at least 240 amino acids that has at least 90% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22, or a second nucleotide sequence comprising the complement of the first nucleotide sequence.

2. The isolated polynucleotide of Claim 1, wherein the first nucleotide sequence consists of a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19 and 21 that codes for the polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22.

3. The isolated polynucleotide of Claim 1 wherein the nucleotide sequences are DNA.

4. The isolated polynucleotide of Claim 1 wherein the nucleotide sequences are RNA.

5. A chimeric gene comprising the isolated polynucleotide of Claim 1 operably linked to suitable regulatory sequences.

6. An isolated host cell comprising the chimeric gene of Claim 5.

7. A host cell comprising an isolated polynucleotide of Claim 1.

8. The host cell of Claim 7 wherein the host cell is selected from the group consisting of yeast, bacteria, plant, and virus.

9. A virus comprising the isolated polynucleotide of Claim 1.

10. A polypeptide of at least 240 amino acids that has at least 90% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22.

11. A method of selecting an isolated polynucleotide that affects the level of expression of a carbon catabolite repression polypeptide in a plant cell, the method comprising the steps of:

(a) constructing an isolated polynucleotide comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from an isolated polynucleotide of Claim 1;

(b) introducing the isolated polynucleotide into a plant cell;

(c) measuring the level of a polypeptide in the plant cell containing the polynucleotide to provide a positive selection means.

12. The method of Claim 11 wherein the isolated polynucleotide consists of a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19 and 21 that codes for the polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22.

13. A method of selecting an isolated polynucleotide that affects the level of expression of a carbon catabolite repression polypeptide in a plant cell, the method comprising the steps of:

(a) constructing an isolated polynucleotide of Claim 1;

(b) introducing the isolated polynucleotide into a plant cell; and

(c) measuring the level of polypeptide in the plant cell containing the polynucleotide to provide a positive selection means.

14. A method of obtaining a nucleic acid fragment encoding a carbon catabolite repression polypeptide comprising the steps of:

(a) synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from a sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, and the complement of such nucleotide sequences; and

(b) amplifying a nucleic acid sequence using the oligonucleotide primer.

15. A method of obtaining a nucleic acid fragment encoding a carbon catabolite repression polypeptide comprising the steps of:

(a) probing a cDNA or genomic library with an isolated polynucleotide comprising at least one of 30 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, and the complement of such nucleotide sequences;

(b) identifying a DNA clone that hybridizes with the isolated polynucleotide;

(c) isolating the identified DNA clone; and

(d) sequencing the cDNA or genomic fragment that comprises the isolated DNA clone.

16. A composition comprising the isolated polynucleotide of Claim 1.

17. An isolated polynucleotide comprising the nucleotide sequence having at least one of 30 contiguous nucleotides derived from a nucleic acid sequence selected from the group consisting of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, and the complement of such sequences.

18. An expression cassette comprising an isolated polynucleotide of Claim 1 operably linked to a promoter.

19. A method for positive selection of a transformed cell comprising:

(a) transforming a host cell with the chimeric gene of Claim 5 or an expression cassette of Claim 18; and

(b) growing the transformed host cell under conditions which allow expression of the polynucleotide in an amount sufficient to alter expression of glucose repressible genes to provide a positive selection means.

20. The method of Claim 19 wherein the host is a plant cell.

21 The method of Claim 19 wherein the plant cell is a dicot or a monocot.

22. An isolated polynucleotide comprising a first nucleotide sequence encoding a polypeptide of at least 77 amino acids that has at least 85% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of SEQ ID NOs:24, 26 and 28, or a second nucleotide sequence comprising the complement of the first nucleotide sequence.

23. A polypeptide of at least 77 amino acids that has at least 85% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of SEQ ID NOs:24, 26 and 28, or a second nucleotide sequence comprising the complement of the first nucleotide sequence.

SEQUENCE LISTING

<110> E. I. du Pont de Nemours and Company
Pioneer Hi-Bred International, Inc.

<120> Plant Carbon Catabolite Repression Proteins

<130> BB1315

<140>
<141>

<150> 60/112,563
<151> 1998-12-16

<160> 28

<170> Microsoft Office 97

<210> 1
<211> 1707
<212> DNA
<213> Zea mays

<400> 1
gcacgaggag cagctcccct gcccctcgca gcggctactc tacaggtcta gcgactcttt 60
cgccatccat agagggagga ggcgcggcgg agatggtggg cggtggcggc ggcgggccgc 120
tgcggcgggt gggcaagtac gaggtgggac gcaccatcgg ggaaggcacc ttcgccaagg 180
tcaagttcgc gcagaacacc gagaccgggg agagcgtcgc catgaaggtg ctcgaccgct 240
cctccatcct caagaacaag atggccgaac agattaagag agaaatatcc ataatgaagc 300
ttgtcaggca tcccaatgtc gttaggctac acgaggtttt ggcaagccgg aagaagatat 360
ttataattct ggagttcatc actggcggcg agctattcga taaaattatt cgtcatggga 420
gactcagtga agcagatgcc cgcagatact ttcagcagct tattgatggt gttgattttt 480
gtcacaagaa aggagtctac catcgagact taaagcctga aaatctcctc cttgattccc 540
aaggcaatct taaaatttca gactttggac tgagtgcatg gcctgctcag ggatctttcc 600
ttctacgtac aacctgtggg actccgaact atgttgcccc agaggttctc agtcataagg 660
gatataatgg cgcacttgct gatacatggt catgtggagt aattttatat gtattgttag 720
caggttatct tccatttgat gaagtggacc tgactaccct ttatggaaag attgagagtg 780
cagaatattc attcccagcc tggttttcag gtggtgcaaa atcactgatt cgtagaattc 840
ttgacccaaa tccagagaca cgaatcagga tagaagagat cagaagcgat gaatggtttc 900
agaaaaatta tgaacctatc aaagaaatag aaaatgagga agtcaatctc gatgatgtca 960
atgcagcttt tgatgatcct gaggacgata atgaggatgc tttcgaagat gaaacaggtc 1020
ctctaacact taatgcattc gacctaatca ttctttcgca aggattgaat cttgcagcgc 1080
tctttgaccg cagacaggac tgtgacaagc ttcaaaatag attcttatca cgcaatccag 1140
caaaggttat cttgtcaagc atggaggttg ttgcgcaatc aatgggattt aagacacaca 1200
ttcgcaatta taagatgagg gtggaaggtc taaatgccga caagactagt catctctcag 1260
ttatggttga agttttcgaa gttgctccat caatcttcat ggtagaacta caaagagcag 1320
caggagatac ttcagagtat aacacgttcg taaataacta ctgcggcaaa ttagatgata 1380
tcatctggaa atttccaact gagaagggaa aatcaaggat accccggtta tcaaagtctc 1440
attcataatc tagtgctgtg tgttatcaat tactatggac ttaagtggat atatttgcag 1500
agtgtatcgc aattggatta tgatattcta agtgtagctc ttcgagtagt atatgtttgt 1560
atattcagag tccccatgta caaagaggcc gttttaattg aggtttccgt acctcgattc 1620
cctgtgcaaa gttttcatcg atgttttaga ctgcacatct gtatattttg tcgagagcat 1680
cagtttctta aaaaaaaaaa aaaaaaa 1707

<210> 2
<211> 481
<212> PRT
<213> Zea mays

<400> 2

```
Thr Arg Ser Ser Ser Pro Ala Pro Arg Ser Gly Tyr Ser Thr Gly Leu
  1               5                  10                  15

Ala Thr Leu Ser Pro Ser Ile Glu Gly Gly Gly Ala Ala Glu Met Val
             20                  25                  30

Gly Gly Gly Gly Gly Gly Pro Leu Arg Arg Val Gly Lys Tyr Glu Val
         35                  40                  45

Gly Arg Thr Ile Gly Glu Gly Thr Phe Ala Lys Val Lys Phe Ala Gln
     50                  55                  60

Asn Thr Glu Thr Gly Glu Ser Val Ala Met Lys Val Leu Asp Arg Ser
 65                  70                  75                  80

Ser Ile Leu Lys Asn Lys Met Ala Glu Gln Ile Lys Arg Glu Ile Ser
                 85                  90                  95

Ile Met Lys Leu Val Arg His Pro Asn Val Val Arg Leu His Glu Val
            100                 105                 110

Leu Ala Ser Arg Lys Lys Ile Phe Ile Ile Leu Glu Phe Ile Thr Gly
        115                 120                 125

Gly Glu Leu Phe Asp Lys Ile Ile Arg His Gly Arg Leu Ser Glu Ala
    130                 135                 140

Asp Ala Arg Arg Tyr Phe Gln Gln Leu Ile Asp Gly Val Asp Phe Cys
145                 150                 155                 160

His Lys Lys Gly Val Tyr His Arg Asp Leu Lys Pro Glu Asn Leu Leu
                165                 170                 175

Leu Asp Ser Gln Gly Asn Leu Lys Ile Ser Asp Phe Gly Leu Ser Ala
            180                 185                 190

Trp Pro Ala Gln Gly Ser Phe Leu Leu Arg Thr Thr Cys Gly Thr Pro
        195                 200                 205

Asn Tyr Val Ala Pro Glu Val Leu Ser His Lys Gly Tyr Asn Gly Ala
    210                 215                 220

Leu Ala Asp Thr Trp Ser Cys Gly Val Ile Leu Tyr Val Leu Leu Ala
225                 230                 235                 240

Gly Tyr Leu Pro Phe Asp Glu Val Asp Leu Thr Thr Leu Tyr Gly Lys
                245                 250                 255

Ile Glu Ser Ala Glu Tyr Ser Phe Pro Ala Trp Phe Ser Gly Gly Ala
            260                 265                 270

Lys Ser Leu Ile Arg Arg Ile Leu Asp Pro Asn Pro Glu Thr Arg Ile
        275                 280                 285

Arg Ile Glu Glu Ile Arg Ser Asp Glu Trp Phe Gln Lys Asn Tyr Glu
    290                 295                 300

Pro Ile Lys Glu Ile Glu Asn Glu Glu Val Asn Leu Asp Asp Val Asn
305                 310                 315                 320
```

```
Ala Ala Phe Asp Asp Pro Glu Asp Asp Asn Glu Asp Ala Phe Glu Asp
                325                 330                 335

Glu Thr Gly Pro Leu Thr Leu Asn Ala Phe Asp Leu Ile Ile Leu Ser
            340                 345                 350

Gln Gly Leu Asn Leu Ala Ala Leu Phe Asp Arg Arg Gln Asp Cys Asp
        355                 360                 365

Lys Leu Gln Asn Arg Phe Leu Ser Arg Asn Pro Ala Lys Val Ile Leu
    370                 375                 380

Ser Ser Met Glu Val Val Ala Gln Ser Met Gly Phe Lys Thr His Ile
385                 390                 395                 400

Arg Asn Tyr Lys Met Arg Val Glu Gly Leu Asn Ala Asp Lys Thr Ser
                405                 410                 415

His Leu Ser Val Met Val Glu Val Phe Glu Val Ala Pro Ser Ile Phe
            420                 425                 430

Met Val Glu Leu Gln Arg Ala Ala Gly Asp Thr Ser Glu Tyr Asn Thr
        435                 440                 445

Phe Val Asn Asn Tyr Cys Gly Lys Leu Asp Asp Ile Ile Trp Lys Phe
    450                 455                 460

Pro Thr Glu Lys Gly Lys Ser Arg Ile Pro Arg Leu Ser Lys Ser His
465                 470                 475                 480

Ser
```

<210> 3
<211> 1948
<212> DNA
<213> Zea mays

<400> 3

```
gtcgacccac gcgtccggac caaagccggg catacgcccc caagtccaaa agccctctcc   60
gccccgctct cccactcgta ggtgttctcc ccgtctccgc ccgcactcgc tcgtccgcgc  120
gcaggaaggt tgacctgtcg agggccggcg aacccggtaa gtaagagtga aaatggatgg  180
aagtagtaaa gggagtgggc attctgaagc attaaggaac tacaacctgg gaagaacttt  240
aggtatcggt acatttggaa aagtgaagat tgcagagcat aagcttactg gacatagggt  300
tgctataaag atcatcaact gccgccaaat gagaaatatg gaaatggaag agaaagcaaa  360
gagagaattc aagatattga agttgttcat tcacccccat atcattcggc tttatgaggt  420
catatacaca cctacagata tatatgttgt gatggaatat tgtaagtatg gcgagttatt  480
tgattacatt gttgagaaag gcagattaca ggaagatgaa gctcgtcgaa tcttccagca  540
gatcatatct ggcgtcgaat actgccatag aaacatggtt gtccaccgtg acctaaagcc  600
ggaaaacttg ttacttgatt caaagtataa tgtaaaactt gcggattttg gtctgagcaa  660
tgtcatgcat gatggccatt ttctgaagac tagctgtggg agtccgaact atgctgctcc  720
agaggtaata tctggtaaac tatatgctgg acctgaggtc gatgtatgga gttgtggggt  780
gattctttat gctcttcttt gtggaactct tccatttgat gatgagaata ttcccaatct  840
gttcaaaaaa attaagggag gtatctacac acttccaagt catttgtctg ctttggccag  900
ggatttgatc ccacgaatgc ttgttgttga gcctatgaag agaatcacaa ttagggaaat  960
tcgggagcat caatggttcc agattcgcct tccacgttac ttggcagtgc ctccaccaga 1020
tacgacacaa caagccaaaa tgattgatga agatacactt cgagatgttg ttaatatggg 1080
atttaacaag aaccatgtgt gtgaatcact gtgcagcaga cttcaaaatg aggcaactgt 1140
tgcatattat ttactattgg acaatcggtt tagagcaact agtggctatc ttggggcaga 1200
```

```
ttatcaagaa tcaatggaca ggaatttaaa tcagctggcg tcatctgaat catctagttc 1260
tggtacgagg aattatgttc caggaagcag tgatcctcat agcagtggtt tgcggccata 1320
ttatcctgtt gaaagaaaat gggcgcttgg acttcagtct cgggcccacc ctcgtgaaat 1380
aatggttgag gtcttaaaag cacttcaaga attaaacgtc agatggaaga agaatgggca 1440
ctacaacgtg aaatgcagat ggtgcccagg gtttcctgaa gttaatgaca cgttagatgc 1500
cagcaacagc tttcttggtg actctaccat catggataat gatgatgcta atgggaggct 1560
acctactgtg atcaagtttg aattccagct ttacaagacg aaggacgaca agtacctctt 1620
agatatgcag agagttactg gacctcagct gctcttcctt gacttctgtg cggccttcct 1680
taccaagctt agggttctat agtggtctac catgtgcaaa ttttcactgt ggtgatgaat 1740
aaccgaaagc atgtaaatag gaaccttgtt ctcgtctttt ggacaacgaa catgtttgag 1800
tgactggtct tgtgttgagc gcgtaaaggt catgtatact taggttagta ctattttcgt 1860
tcttaaatat ttgtcgtctg ctagtgatag ttcatttttg aactaaaacg ttacgaataa 1920
aaaaagagta aaaaaaaaaa aaaaaaaa                                    1948
```

<210> 4
<211> 509
<212> PRT
<213> Zea mays

<400> 4

```
Met Asp Gly Ser Ser Lys Gly Ser Gly His Ser Glu Ala Leu Arg Asn
  1               5                  10                  15

Tyr Asn Leu Gly Arg Thr Leu Gly Ile Gly Thr Phe Gly Lys Val Lys
             20                  25                  30

Ile Ala Glu His Lys Leu Thr Gly His Arg Val Ala Ile Lys Ile Ile
         35                  40                  45

Asn Cys Arg Gln Met Arg Asn Met Glu Met Glu Glu Lys Ala Lys Arg
     50                  55                  60

Glu Phe Lys Ile Leu Lys Leu Phe Ile His Pro His Ile Ile Arg Leu
 65                  70                  75                  80

Tyr Glu Val Ile Tyr Thr Pro Thr Asp Ile Tyr Val Val Met Glu Tyr
                 85                  90                  95

Cys Lys Tyr Gly Glu Leu Phe Asp Tyr Ile Val Glu Lys Gly Arg Leu
            100                 105                 110

Gln Glu Asp Glu Ala Arg Arg Ile Phe Gln Gln Ile Ile Ser Gly Val
        115                 120                 125

Glu Tyr Cys His Arg Asn Met Val Val His Arg Asp Leu Lys Pro Glu
    130                 135                 140

Asn Leu Leu Leu Asp Ser Lys Tyr Asn Val Lys Leu Ala Asp Phe Gly
145                 150                 155                 160

Leu Ser Asn Val Met His Asp Gly His Phe Leu Lys Thr Ser Cys Gly
                165                 170                 175

Ser Pro Asn Tyr Ala Ala Pro Glu Val Ile Ser Gly Lys Leu Tyr Ala
            180                 185                 190

Gly Pro Glu Val Asp Val Trp Ser Cys Gly Val Ile Leu Tyr Ala Leu
        195                 200                 205
```

```
Leu Cys Gly Thr Leu Pro Phe Asp Asp Glu Asn Ile Pro Asn Leu Phe
    210                 215                 220

Lys Lys Ile Lys Gly Gly Ile Tyr Thr Leu Pro Ser His Leu Ser Ala
225                 230                 235                 240

Leu Ala Arg Asp Leu Ile Pro Arg Met Leu Val Val Glu Pro Met Lys
                245                 250                 255

Arg Ile Thr Ile Arg Glu Ile Arg Glu His Gln Trp Phe Gln Ile Arg
            260                 265                 270

Leu Pro Arg Tyr Leu Ala Val Pro Pro Asp Thr Thr Gln Gln Ala
        275                 280                 285

Lys Met Ile Asp Glu Asp Thr Leu Arg Asp Val Val Asn Met Gly Phe
    290                 295                 300

Asn Lys Asn His Val Cys Glu Ser Leu Cys Ser Arg Leu Gln Asn Glu
305                 310                 315                 320

Ala Thr Val Ala Tyr Tyr Leu Leu Leu Asp Asn Arg Phe Arg Ala Thr
                325                 330                 335

Ser Gly Tyr Leu Gly Ala Asp Tyr Gln Glu Ser Met Asp Arg Asn Leu
            340                 345                 350

Asn Gln Leu Ala Ser Ser Glu Ser Ser Ser Ser Gly Thr Arg Asn Tyr
        355                 360                 365

Val Pro Gly Ser Ser Asp Pro His Ser Ser Gly Leu Arg Pro Tyr Tyr
    370                 375                 380

Pro Val Glu Arg Lys Trp Ala Leu Gly Leu Gln Ser Arg Ala His Pro
385                 390                 395                 400

Arg Glu Ile Met Val Glu Val Leu Lys Ala Leu Gln Glu Leu Asn Val
                405                 410                 415

Arg Trp Lys Lys Asn Gly His Tyr Asn Val Lys Cys Arg Trp Cys Pro
            420                 425                 430

Gly Phe Pro Glu Val Asn Asp Thr Leu Asp Ala Ser Asn Ser Phe Leu
        435                 440                 445

Gly Asp Ser Thr Ile Met Asp Asn Asp Asp Ala Asn Gly Arg Leu Pro
    450                 455                 460

Thr Val Ile Lys Phe Glu Phe Gln Leu Tyr Lys Thr Lys Asp Asp Lys
465                 470                 475                 480

Tyr Leu Leu Asp Met Gln Arg Val Thr Gly Pro Gln Leu Leu Phe Leu
                485                 490                 495

Asp Phe Cys Ala Ala Phe Leu Thr Lys Leu Arg Val Leu
            500                 505
```

<210> 5
<211> 2107

<212> DNA
<213> Zea mays

<400> 5
gtcgacccac gcgtccgccc atgtgagcct cgcttttctc ctctctctct ccccggcatt 60
tctctctctg gcgtcactcg cgtgtcaggc gcattctgcc tctttctttc tccccctccc 120
ctcccctccc ctgcgggctt ctctcacgag ctccccccggc ctcgccgccg ccgcctgctc 180
cggcgggcgg gacccagctc gcagcgtctc ccctgcggcg aggtacacga tggagggagc 240
gggaagagat gccaaccctt tgagcggtta cagaattggc aaaaccctgg gaattgggtc 300
gttcggtaaa gtgaagatcg cggaacatat attgactggt cataaggtgg cgatcaagat 360
tctcaatcgc aagaagatca gaagcatgga tatggaagag aaagttaaga gagaaatcaa 420
gatactgaga ttatttatgc atcctcatat catacgcctt tatgaggtga tagatacacc 480
tgctgatatc tgtgttgtta tggagtatgt taaatctgga gagttgtttg attacatcgt 540
tgagaaggga aggctacacg aagaggaagc ccgacacttt tttcagcaga tcatatctgg 600
tgttgaatat tgccatagga acatggttgc tcaccgtgat ttaaagccag agaatcttct 660
tttggattca aaatgcaatg ttaagattgc cgattttggc ttaagtaata ttatgcgtga 720
tggtcacttt cttaagacga gttgtggtag cccgaattat gcagcacctg aggtcatatc 780
tggtaaacta tatgctggtc ctgaagttga cgtctggagc tgtggagtta ttctttatgc 840
tcttctttgt ggcactctcc catttgacga tgagaatatt ccaaaccttt tcaagaaaat 900
aaagggtgga atatataccc ttcctagtca tttgtcacct tcagcgaggg acttgattcc 960
cagaatgctg gttgttgatc caatgaaaag gattacaata cgtgaaatcc gtgaacatgt 1020
gtggttcaag atccgacttc cgcgctattt ggctgtgccg cctccagaca ctgctcaaca 1080
agttaaaaag gtcgacgagg aaactcttaa tgatgttatt aagatgggtt ttgacaagaa 1140
tcagctaatt gaatctctgc aaaacagatt gcagaatgag gcaacagttg cctattattt 1200
actcttggac aataggcttc gtacaaccag tggttatctt ggatctgagt ttcaagaatc 1260
tatggactca tctttgtctc aagtaatcgc tgaaacacca acttcagcaa ctgaacttcg 1320
tcagcatggg ttttcagaat ctccaggttc tggcttgagg cagcattttg cagctgaaag 1380
gaaatgggcc cttggtcttc agtctcgagc acatccacga gaaataataa gtgaagtgct 1440
taaagctctg caagaactga atgtttactg gaaaaagatt ggacactaca acatgaaatg 1500
cagatggagt cctggctgcc ttgagagtat gatgcataac agtgatagct tcagtgcgga 1560
gtctgctata attgaaactg atgttttcat ggagaaatca accccgacag tgaagtttga 1620
gattcagctt tacaaaacga gggatgagaa gtaccttctt gacctgcaaa gggtcagtgg 1680
atcacatctt ctctttctgg acttgtgttc cgcctttcta actcagctga gagttctttg 1740
agcctgacat gattggcttc ccaagcactc cttgatgcac ccagtggagt ttcaagttac 1800
atttgagatg tacataatga ccatatataa tcctccagta gattatattt tctggagcat 1860
gtaatgtaga cgttagaggc ttgtcctggt ttgccaagat ggttgttacc taaggtccta 1920
acaagaggcc gcaggcgagc ccagtccaat tagatgatgt gaaaaatttt tggccttttt 1980
ttagtctttg gacatgttat caaattaagg aggtcgtgct agcacgatac ccaaactaag 2040
aaatacccca ttgaaaactg aaaacaaagc ttgctcgttc aaaaaaaaaa aaaaaaaaaa 2100
aaaaaaa 2107

<210> 6
<211> 579
<212> PRT
<213> Zea mays

<400> 6
Ser Thr His Ala Ser Ala His Val Ser Leu Ala Phe Leu Leu Ser Leu
1 5 10 15

Ser Pro Ala Phe Leu Ser Leu Ala Ser Leu Ala Cys Gln Ala His Ser
20 25 30

Ala Ser Phe Phe Leu Pro Leu Pro Ser Pro Pro Leu Arg Ala Ser Leu
35 40 45

Thr Ser Ser Pro Gly Leu Ala Ala Ala Ala Cys Ser Gly Gly Arg Asp
50 55 60

```
Pro Ala Arg Ser Val Ser Pro Ala Ala Arg Tyr Thr Met Glu Gly Ala
 65                  70                  75                  80

Gly Arg Asp Ala Asn Pro Leu Ser Gly Tyr Arg Ile Gly Lys Thr Leu
                 85                  90                  95

Gly Ile Gly Ser Phe Gly Lys Val Lys Ile Ala Glu His Ile Leu Thr
            100                 105                 110

Gly His Lys Val Ala Ile Lys Ile Leu Asn Arg Lys Lys Ile Arg Ser
        115                 120                 125

Met Asp Met Glu Glu Lys Val Lys Arg Glu Ile Lys Ile Leu Arg Leu
    130                 135                 140

Phe Met His Pro His Ile Ile Arg Leu Tyr Glu Val Ile Asp Thr Pro
145                 150                 155                 160

Ala Asp Ile Cys Val Val Met Glu Tyr Val Lys Ser Gly Glu Leu Phe
                165                 170                 175

Asp Tyr Ile Val Glu Lys Gly Arg Leu His Glu Glu Glu Ala Arg His
            180                 185                 190

Phe Phe Gln Gln Ile Ile Ser Gly Val Glu Tyr Cys His Arg Asn Met
        195                 200                 205

Val Ala His Arg Asp Leu Lys Pro Glu Asn Leu Leu Leu Asp Ser Lys
    210                 215                 220

Cys Asn Val Lys Ile Ala Asp Phe Gly Leu Ser Asn Ile Met Arg Asp
225                 230                 235                 240

Gly His Phe Leu Lys Thr Ser Cys Gly Ser Pro Asn Tyr Ala Ala Pro
                245                 250                 255

Glu Val Ile Ser Gly Lys Leu Tyr Ala Gly Pro Glu Val Asp Val Trp
            260                 265                 270

Ser Cys Gly Val Ile Leu Tyr Ala Leu Leu Cys Gly Thr Leu Pro Phe
        275                 280                 285

Asp Asp Glu Asn Ile Pro Asn Leu Phe Lys Lys Ile Lys Gly Gly Ile
    290                 295                 300

Tyr Thr Leu Pro Ser His Leu Ser Pro Ser Ala Arg Asp Leu Ile Pro
305                 310                 315                 320

Arg Met Leu Val Val Asp Pro Met Lys Arg Ile Thr Ile Arg Glu Ile
                325                 330                 335

Arg Glu His Val Trp Phe Lys Ile Arg Leu Pro Arg Tyr Leu Ala Val
            340                 345                 350

Pro Pro Pro Asp Thr Ala Gln Gln Val Lys Lys Val Asp Glu Glu Thr
        355                 360                 365

Leu Asn Asp Val Ile Lys Met Gly Phe Asp Lys Asn Gln Leu Ile Glu
    370                 375                 380
```

```
Ser Leu Gln Asn Arg Leu Gln Asn Glu Ala Thr Val Ala Tyr Tyr Leu
385                 390                 395                 400

Leu Leu Asp Asn Arg Leu Arg Thr Thr Ser Gly Tyr Leu Gly Ser Glu
                405                 410                 415

Phe Gln Glu Ser Met Asp Ser Ser Leu Ser Gln Val Ile Ala Glu Thr
            420                 425                 430

Pro Thr Ser Ala Thr Glu Leu Arg Gln His Gly Phe Ser Glu Ser Pro
        435                 440                 445

Gly Ser Gly Leu Arg Gln His Phe Ala Ala Glu Arg Lys Trp Ala Leu
    450                 455                 460

Gly Leu Gln Ser Arg Ala His Pro Arg Glu Ile Ile Ser Glu Val Leu
465                 470                 475                 480

Lys Ala Leu Gln Glu Leu Asn Val Tyr Trp Lys Lys Ile Gly His Tyr
                485                 490                 495

Asn Met Lys Cys Arg Trp Ser Pro Gly Cys Leu Glu Ser Met Met His
            500                 505                 510

Asn Ser Asp Ser Phe Ser Ala Glu Ser Ala Ile Ile Glu Thr Asp Val
        515                 520                 525

Phe Met Glu Lys Ser Thr Pro Thr Val Lys Phe Glu Ile Gln Leu Tyr
    530                 535                 540

Lys Thr Arg Asp Glu Lys Tyr Leu Leu Asp Leu Gln Arg Val Ser Gly
545                 550                 555                 560

Ser His Leu Leu Phe Leu Asp Leu Cys Ser Ala Phe Leu Thr Gln Leu
                565                 570                 575

Arg Val Leu
```

<210> 7
<211> 530
<212> DNA
<213> Oryza sativa

<220>
<221> unsure
<222> (466)

<220>
<221> unsure
<222> (503)

<220>
<221> unsure
<222> (507)

<400> 7

```
gagggtaggg gactacgtgc tggtgcggca gatcgggtcg ggggcgtacg cgcgggtctg  60
gctcgggaag caccgcacgc ggggcacgga ggtggccttg aaggagatcg ccgtggagcg 120
gcttagcagc aagctccgcg agagcctcct ctccgaggtc gacatcctcc ggcgcatccg 180
```

tcatcccaac gtcatcgccc tccacgagtc catcagggat ggtgggaaaa tatatcttgt 240
attagagtac tgtcgaggtg gtgacttaca ctcatacctt cagcagcata aaagggtttc 300
tgaaacagtt gctaagcatt tcatccagca gctagcatct ggtctgcaga tgctgcgtga 360
aaacaacgtg gttcatcgag atctaaaaac cacagaaatt cttctaattg caaataatga 420
aaatctcccc ttgaagattg cggactttgg atttgcaaat ttttanaacc ttcttytctg 480
ggctgaaaac actttgcggt tcncgcntta aaagggtcca aaagtcatca 530

<210> 8
<211> 151
<212> PRT
<213> Oryza sativa

<400> 8
Val Gly Asp Tyr Val Leu Val Arg Gln Ile Gly Ser Gly Ala Tyr Ala
1 5 10 15

Arg Val Trp Leu Gly Lys His Arg Thr Arg Gly Thr Glu Val Ala Leu
20 25 30

Lys Glu Ile Ala Val Glu Arg Leu Ser Ser Lys Leu Arg Glu Ser Leu
35 40 45

Leu Ser Glu Val Asp Ile Leu Arg Arg Ile Arg His Pro Asn Val Ile
50 55 60

Ala Leu His Glu Ser Ile Arg Asp Gly Gly Lys Ile Tyr Leu Val Leu
65 70 75 80

Glu Tyr Cys Arg Gly Gly Asp Leu His Ser Tyr Leu Gln Gln His Lys
85 90 95

Arg Val Ser Glu Thr Val Ala Lys His Phe Ile Gln Gln Leu Ala Ser
100 105 110

Gly Leu Gln Met Leu Arg Glu Asn Asn Val Val His Arg Asp Leu Lys
115 120 125

Thr Thr Glu Ile Leu Leu Ile Ala Asn Asn Glu Asn Leu Pro Leu Lys
130 135 140

Ile Ala Asp Phe Gly Phe Ala
145 150

<210> 9
<211> 1848
<212> DNA
<213> Glycine max

<400> 9
ggaatgcaga gaaatgagaa ttgggcattg gctcagtcct tctgaggcag gtggcgctgc 60
caaacgcaga acctccggca gaaattgcca ccgcaaacac gtggaagacg ggtagtagta 120
ctccatctcc cattctctca tattcccttt actaattaaa gatcatgaac agcgagagac 180
agtcgacgac gacgacgttg ctgcacggca agtacgagct aggtcgtgtg ctggggcacg 240
gaagcttcgc caaggtctac cacgcgcgga acctgaagac ggggcagcac gtggctatga 300
aggtcgttgg aaaagaaaag gtgatcaagg tcggaatgat ggagcaggtc aagagggaga 360
tctcggtcat gaagatggtc aagcacccaa acatcgtcga gctccacgaa gtcatggcca 420
gtaagtccaa gatctacatc tccatcgaac tcgtccgcgg cggagagctc ttcaacaagg 480
tctccaaggg acgcttgaag gaggacctgg ccagactcta cttccagcag ttaatctctg 540
ccgtcgattt ctgccacagc cgcggcgtct accaccgtga cctcaagccg gagaatctcc 600
tcctggacga acacggcaac ctcaaggtct ccgacttcgg actcaccgcc ttctccgacc 660

```
accttaaaga ggacgggctg ttacacacca cgtgcggcac gcctgcgtac gtgtcaccgg  720
aagttattgc aaagaaaggc tatgacggtg ccaaggctga tatatggtca tgcggagtaa  780
tcctctacgt tctcctagct gggtttttac cctttcagga tgataatttg gttgccatgt  840
ataaaaaaat ccatagaggg gacttcaagt gcccaccgtg gttctctctc gatgcgagaa  900
agcttgttac gaaactgctt gatccgaatc cgaatacgag gattagtatt tctaaagtga  960
tggagtcttc ttggtttaag aaacaggtgc cgaggaaggt ggaggaggtg gttgagaagg 1020
tcgatttgga ggagaagatt gagaatcagg agacgatgaa tgctttccac ataatctcct 1080
tgtcggaggg gttcaatttg tcgccgttgt tcgaggagaa gaggaaggag gagatgaggt 1140
tcgccaccgc ggggacgccg agcagcgtca tttcgcgcct ggaggaggtg gccaaggcgg 1200
ggaagtttga cgtgaagagt agcgagacga aagtgaggct tcagggacag gagcgtggga 1260
ggaaggggaa gctagcgatt gcggcggata tctacgccgt gacgccgtcg tttatggtgg 1320
tggaggtcaa gaaggacaat ggggatacgt tggagtataa ccagttctgc agcaaacaac 1380
ttcgtcctgc gcttaaggat atcttctgga attctgcacc tgccagtgct tgaatggtat 1440
attatatata tgatgtctga ttattgtttt acaacaccat ttgtcgacct tgttatttta 1500
ttttattttt gtttaggaac tttgagtttt gttgataaga tctgtttgtg ttttgttgtt 1560
taaattgtta gtttggtact tggggactgg ttagtgggta cttactctct ctgtcgccca 1620
gaattgttgt taagagatgc acacgttgtg gatgattctg tcatgtttgt gtttcaatta 1680
tgagtgtgta atttggagtg cttgtggggc ttactgcgac tttgttcaaa gttacttttt 1740
acctttttct tttctttatt ttattccttc tagattatgt cctctgtgtt cacgggtcac 1800
gcctgtacct aatacactag gaattcgact aaaaaaaaaa aaaaaaaa              1848
```

<210> 10
<211> 422
<212> PRT
<213> Glycine max

<400> 10

```
Met Asn Ser Glu Arg Gln Ser Thr Thr Thr Thr Leu Leu His Gly Lys
  1               5                  10                  15

Tyr Glu Leu Gly Arg Val Leu Gly His Gly Ser Phe Ala Lys Val Tyr
             20                  25                  30

His Ala Arg Asn Leu Lys Thr Gly Gln His Val Ala Met Lys Val Val
         35                  40                  45

Gly Lys Glu Lys Val Ile Lys Val Gly Met Met Glu Gln Val Lys Arg
     50                  55                  60

Glu Ile Ser Val Met Lys Met Val Lys His Pro Asn Ile Val Glu Leu
 65                  70                  75                  80

His Glu Val Met Ala Ser Lys Ser Lys Ile Tyr Ile Ser Ile Glu Leu
                 85                  90                  95

Val Arg Gly Gly Glu Leu Phe Asn Lys Val Ser Lys Gly Arg Leu Lys
            100                 105                 110

Glu Asp Leu Ala Arg Leu Tyr Phe Gln Gln Leu Ile Ser Ala Val Asp
        115                 120                 125

Phe Cys His Ser Arg Gly Val Tyr His Arg Asp Leu Lys Pro Glu Asn
    130                 135                 140

Leu Leu Leu Asp Glu His Gly Asn Leu Lys Val Ser Asp Phe Gly Leu
145                 150                 155                 160

Thr Ala Phe Ser Asp His Leu Lys Glu Asp Gly Leu Leu His Thr Thr
                165                 170                 175
```

```
Cys Gly Thr Pro Ala Tyr Val Ser Pro Glu Val Ile Ala Lys Lys Gly
            180                 185                 190

Tyr Asp Gly Ala Lys Ala Asp Ile Trp Ser Cys Gly Val Ile Leu Tyr
        195                 200                 205

Val Leu Leu Ala Gly Phe Leu Pro Phe Gln Asp Asp Asn Leu Val Ala
    210                 215                 220

Met Tyr Lys Lys Ile His Arg Gly Asp Phe Lys Cys Pro Pro Trp Phe
225                 230                 235                 240

Ser Leu Asp Ala Arg Lys Leu Val Thr Lys Leu Leu Asp Pro Asn Pro
                245                 250                 255

Asn Thr Arg Ile Ser Ile Ser Lys Val Met Glu Ser Ser Trp Phe Lys
            260                 265                 270

Lys Gln Val Pro Arg Lys Val Glu Glu Val Val Glu Lys Val Asp Leu
        275                 280                 285

Glu Glu Lys Ile Glu Asn Gln Glu Thr Met Asn Ala Phe His Ile Ile
    290                 295                 300

Ser Leu Ser Glu Gly Phe Asn Leu Ser Pro Leu Phe Glu Glu Lys Arg
305                 310                 315                 320

Lys Glu Glu Met Arg Phe Ala Thr Ala Gly Thr Pro Ser Ser Val Ile
                325                 330                 335

Ser Arg Leu Glu Glu Val Ala Lys Ala Gly Lys Phe Asp Val Lys Ser
            340                 345                 350

Ser Glu Thr Lys Val Arg Leu Gln Gly Gln Glu Arg Gly Arg Lys Gly
        355                 360                 365

Lys Leu Ala Ile Ala Ala Asp Ile Tyr Ala Val Thr Pro Ser Phe Met
    370                 375                 380

Val Val Glu Val Lys Lys Asp Asn Gly Asp Thr Leu Glu Tyr Asn Gln
385                 390                 395                 400

Phe Cys Ser Lys Gln Leu Arg Pro Ala Leu Lys Asp Ile Phe Trp Asn
                405                 410                 415

Ser Ala Pro Ala Ser Ala
            420
```

<210> 11
<211> 2123
<212> DNA
<213> Glycine max

<400> 11

```
cgcgatgctt tgttcctctt ttgttcttcc ccttcccaat cttcaacttt agggttcctt   60
taatctcgtt attcctcttc actcccaagc tccgttcact ccaacaacac tccgatttag  120
aaatggatgg accagctggc cgaggtggtg ctggcctgga catgtttcta ccaaattata  180
aattgggaaa aacactcggg attggatctt ttggcaaggt gaaaattgca gaacatgtgt  240
tgactggcca taaggttgcg atcaagatcc ttaaccgacg caagataaag aacatggaaa  300
tggaagaaaa agtgagaaga gaaatcaaaa ttttaagatt gttcatgcat cctcacatta  360
```

```
ttcgacttta tgaagtcata gaaactccaa ctgacatata tgttgtcatg gagtatgtga  420
agtctggaga gcttttcgat tacatagtag agaagggtag gttgcaggaa gatgaagctc  480
gtaatttttt tcagcagata atctctgggg tggagtactg tcacaggaat atggtggttc  540
atagagattt gaagcctgag aatttacttt tggactccaa atgtaatgtc aagattgctg  600
attttggctt gagcaacatc atgcgtgatg gtcactttct taaaacaagt tgtggaagcc  660
ctaactatgc agctcctgag gttatctctg ggaaattgta tgctggacct gaagtggatg  720
tctggagctg tggtgtaatt ttatatgccc ttctttgtgg cacccttcct tttgatgatg  780
aaaatattcc aaatctcttc aagaaaataa agggtgggat ttacactctt cccagtcatc  840
tatcacccgg tgctagagat ttgataccag ggatgcttgt ggttgaccct atgaggagaa  900
tgaccatacc tgagatccgt caacacccat ggttccaagc tcgacttcca cgttatttag  960
ctgtgccacc accagataca atgcaacagg ccaaaaagat tgatgaggag atccttcagg 1020
aagtggtgaa aatgggattt gacaggaatc aattggttga atctcttggg aacaggatac 1080
aaaatgaggg tactgtggca tactatttgt tattggacaa ccgatttcgt gtttccagtg 1140
gctatcttgg agctgagttt caagagacca tggattccgg ttttaatcaa atgcattcca 1200
gtgaacttgc ttcttcagtt gttggaaacc gctttccagg ctacatggaa tatccaggag 1260
taggatcgag gcaacagttc cctgttgaaa ggaaatgggc ccttgggctt cagtctcgag 1320
cccatcctcg tgaaataatg actgaggttc ttaaagcttt gcaagaatta aatgtttgtt 1380
ggaagaagat tggtcactac aacatgaagt gtaggtgggt tgctggcatt cctggtcacc 1440
acgaaggaat ggttaacaat aatgtgcata gtaatcatta ctttggagat gattccaaca 1500
ttattgagaa tgatgctgtt tctacttcaa atgtggtcaa gtttgaagtg cagctttaca 1560
aaacccggga agaaaagtat ctgcttgatc ttcaaagggt gcagggtcca cagtttcttt 1620
tcttggatct atgtgctgct ttccttgcac agcttcgtgt cctctagagc gaagagctta 1680
ggctcaccag agacataaat atgccttgtg tttcatgtga ataagcttac attgtacata 1740
caagtatctt ttcggtccaa gcttatcctc tgttttctac ccgtttatct gtgggtaaac 1800
attataatga gccaatccca agtatgggat ttgcttcgtt aacaatggtg ctcgactagg 1860
ttttcctttc tcttggtgat tttaatttta tcagtgtaat gaattatatg gggaatgtct 1920
taaaagaatt agccccaact gacagtgata tatgtatggt tatcaaccaa atgctgtaaa 1980
cgactgttgc cattggaagg gatgtctagg ctgcttctaa cttttgcaga ttagtctttg 2040
ttgtcctata aatgtaattt catattactc gtgtatgaat gctatgtaac gactgtgttc 2100
atccgaaaaa aaaaaaaaaa aaa                                        2123
```

<210> 12
<211> 514
<212> PRT
<213> Glycine max

<400> 12

```
Met Asp Gly Pro Ala Gly Arg Gly Gly Ala Gly Leu Asp Met Phe Leu
 1               5                  10                  15

Pro Asn Tyr Lys Leu Gly Lys Thr Leu Gly Ile Gly Ser Phe Gly Lys
            20                  25                  30

Val Lys Ile Ala Glu His Val Leu Thr Gly His Lys Val Ala Ile Lys
        35                  40                  45

Ile Leu Asn Arg Arg Lys Ile Lys Asn Met Glu Met Glu Glu Lys Val
    50                  55                  60

Arg Arg Glu Ile Lys Ile Leu Arg Leu Phe Met His Pro His Ile Ile
65                  70                  75                  80

Arg Leu Tyr Glu Val Ile Glu Thr Pro Thr Asp Ile Tyr Val Val Met
                85                  90                  95

Glu Tyr Val Lys Ser Gly Glu Leu Phe Asp Tyr Ile Val Glu Lys Gly
            100                 105                 110

Arg Leu Gln Glu Asp Glu Ala Arg Asn Phe Phe Gln Gln Ile Ile Ser
        115                 120                 125
```

```
Gly Val Glu Tyr Cys His Arg Asn Met Val Val His Arg Asp Leu Lys
    130                 135                 140

Pro Glu Asn Leu Leu Leu Asp Ser Lys Cys Asn Val Lys Ile Ala Asp
145                 150                 155                 160

Phe Gly Leu Ser Asn Ile Met Arg Asp Gly His Phe Leu Lys Thr Ser
                165                 170                 175

Cys Gly Ser Pro Asn Tyr Ala Ala Pro Glu Val Ile Ser Gly Lys Leu
            180                 185                 190

Tyr Ala Gly Pro Glu Val Asp Val Trp Ser Cys Gly Val Ile Leu Tyr
        195                 200                 205

Ala Leu Leu Cys Gly Thr Leu Pro Phe Asp Asp Glu Asn Ile Pro Asn
    210                 215                 220

Leu Phe Lys Lys Ile Lys Gly Gly Ile Tyr Thr Leu Pro Ser His Leu
225                 230                 235                 240

Ser Pro Gly Ala Arg Asp Leu Ile Pro Gly Met Leu Val Val Asp Pro
                245                 250                 255

Met Arg Arg Met Thr Ile Pro Glu Ile Arg Gln His Pro Trp Phe Gln
            260                 265                 270

Ala Arg Leu Pro Arg Tyr Leu Ala Val Pro Pro Pro Asp Thr Met Gln
        275                 280                 285

Gln Ala Lys Lys Ile Asp Glu Glu Ile Leu Gln Glu Val Val Lys Met
    290                 295                 300

Gly Phe Asp Arg Asn Gln Leu Val Glu Ser Leu Gly Asn Arg Ile Gln
305                 310                 315                 320

Asn Glu Gly Thr Val Ala Tyr Tyr Leu Leu Leu Asp Asn Arg Phe Arg
                325                 330                 335

Val Ser Ser Gly Tyr Leu Gly Ala Glu Phe Gln Glu Thr Met Asp Ser
            340                 345                 350

Gly Phe Asn Gln Met His Ser Ser Glu Leu Ala Ser Ser Val Val Gly
        355                 360                 365

Asn Arg Phe Pro Gly Tyr Met Glu Tyr Pro Gly Val Gly Ser Arg Gln
    370                 375                 380

Gln Phe Pro Val Glu Arg Lys Trp Ala Leu Gly Leu Gln Ser Arg Ala
385                 390                 395                 400

His Pro Arg Glu Ile Met Thr Glu Val Leu Lys Ala Leu Gln Glu Leu
                405                 410                 415

Asn Val Cys Trp Lys Lys Ile Gly His Tyr Asn Met Lys Cys Arg Trp
            420                 425                 430

Val Ala Gly Ile Pro Gly His His Glu Gly Met Val Asn Asn Asn Val
        435                 440                 445
```

```
His Ser Asn His Tyr Phe Gly Asp Asp Ser Asn Ile Ile Glu Asn Asp
    450                 455                 460

Ala Val Ser Thr Ser Asn Val Val Lys Phe Glu Val Gln Leu Tyr Lys
465                 470                 475                 480

Thr Arg Glu Glu Lys Tyr Leu Leu Asp Leu Gln Arg Val Gln Gly Pro
                485                 490                 495

Gln Phe Leu Phe Leu Asp Leu Cys Ala Ala Phe Leu Ala Gln Leu Arg
            500                 505                 510

Val Leu
```

<210> 13
<211> 2040
<212> DNA
<213> Glycine max

<400> 13

```
acttagtgct attcaataat tacgaaaaac caacagcgag tgttgttttt gtttggtcca   60
ggctggacca gggaatgaat agagatgaga ggctttcggt tgaagagaac tgagttaagc  120
ggcgaagcca aacgcagaat cggcggcaga aattgccacc ggaagcacgt cgaagacggg  180
tgaaccaaac acaaaattcc atccaccact tctcttttca atccccaact caaatctgtg  240
aagaaaatct caatttctaa attgtaattc ttagcactac tgctagtagt agtaaaaaaa  300
aacaattaga aagtagtaga gtgttttttt ttctttttt cgtttgtttt acttttaaga  360
tgggtgagaa gagcaacgtt ggcggcgacg cgattaacac gacactgctt cacgggaagt  420
acgagctggg ccggctgtta ggccacggaa ccttcgcgaa ggtgtaccac gcgcgccacc  480
tgaagacggg aaagagcgtg gcgatgaagg tggtgggcaa agagaaggtg gtgaaggtag  540
gcatgatgga gcagatcaag agagagatct cagccatgaa catggtgaag cacccaaaca  600
tcgtgcagct tcacgaggtc atggcgagca agtccaagat ctacatcgcc atggagctcg  660
tccgcggcgg cgagctcttc aacaagatcg ccagaggccg tctccgggag gagatggcca  720
gactttactt ccaacaactc atctctgctg tggacttctg ccacagccgc ggcgtctacc  780
accgtgacct caagccagag aaccttctcc tcgacgacga cggcaacctc aaggtcaccg  840
acttcggcct cagcactttc tccgagcacc tccggcacga cggcctgctg cacactacgt  900
gcggcacgcc ggcctatgtc gcgcccgagg tgattgggaa aagaggctac gacggcgcca  960
aggccgatat ttggtcttgt ggcgtaatcc tctacgtact tcttgctggt tttttaccct 1020
tccaggatga caacctggtg gccttgtaca agaaaattta ccgcggtgac ttcaagtgtc 1080
cgccgtggtt ctcctccgag gcgcggaggc tcatcaccaa gctcctcgac cccaacccga 1140
acacgcgaat cacaatttca aaaatcatgg actcgtcgtg gttcaagaaa cccgtgccga 1200
agaacttgat gggaaagaaa agagaggaat tggatctgga ggagaagatt aaacagcatg 1260
agcaggaggt ttctactacg atgaatgctt ttcacatcat ttcgctctcg gaggggttcg 1320
atttgtcgcc gttgttcgag gaaaagaaga gagaagagaa agagttacgg tttgcgacca 1380
cgcgccccgc gagcagcgtg atttcgagat tggaagattt ggcgaaggcg gtgaagttcg 1440
acgtgaagaa gagcgagact aaggtgaggt tgcagggtca agaaaaaggg cgaaaaggta 1500
aactcgctat tgctgcggac ttgtacgccg tgacgccgtc gtttttggtg gtggaggtta 1560
agaaggacaa cggtgacacc ttggagtata accagttctg tagcaaggag cttcgtccag 1620
ctcttaaaga tatcgtgtgg aggacttctc ctgcagagaa tcctacactc gcttgaagaa 1680
caaccaagac tcccatgcta attgtcgttc ctttgtagtt gtttttgtta attcttgttt 1740
caggactgag atattttagt tatttgcttg cttctttgtt atgttttggc ctctggggtg 1800
tgtcgtcttt taccccgaaa ggggaattgt tgaaaaggcg aatgcagtaa tgaaccgaaa 1860
tggacatcga ggagttgttg tcgatgatga ttcaggtaca tgtgctgcgt gttgttgtgc 1920
tcgtgggtga tgttgttgct atgttggcgt gattgtgaat atgtttgcaa attgattggt 1980
gtggctcatg ggtcattcat ggctgcgttt aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2040
```

<210> 14
<211> 438

<212> PRT
<213> Glycine max

<400> 14

```
Met Gly Glu Lys Ser Asn Val Gly Gly Asp Ala Ile Asn Thr Thr Leu
  1               5                  10                  15

Leu His Gly Lys Tyr Glu Leu Gly Arg Leu Leu Gly His Gly Thr Phe
            20                  25                  30

Ala Lys Val Tyr His Ala Arg His Leu Lys Thr Gly Lys Ser Val Ala
        35                  40                  45

Met Lys Val Val Gly Lys Glu Lys Val Val Lys Val Gly Met Met Glu
    50                  55                  60

Gln Ile Lys Arg Glu Ile Ser Ala Met Asn Met Val Lys His Pro Asn
65                  70                  75                  80

Ile Val Gln Leu His Glu Val Met Ala Ser Lys Ser Lys Ile Tyr Ile
                85                  90                  95

Ala Met Glu Leu Val Arg Gly Gly Glu Leu Phe Asn Lys Ile Ala Arg
            100                 105                 110

Gly Arg Leu Arg Glu Glu Met Ala Arg Leu Tyr Phe Gln Gln Leu Ile
        115                 120                 125

Ser Ala Val Asp Phe Cys His Ser Arg Gly Val Tyr His Arg Asp Leu
    130                 135                 140

Lys Pro Glu Asn Leu Leu Leu Asp Asp Asp Gly Asn Leu Lys Val Thr
145                 150                 155                 160

Asp Phe Gly Leu Ser Thr Phe Ser Glu His Leu Arg His Asp Gly Leu
                165                 170                 175

Leu His Thr Thr Cys Gly Thr Pro Ala Tyr Val Ala Pro Glu Val Ile
            180                 185                 190

Gly Lys Arg Gly Tyr Asp Gly Ala Lys Ala Asp Ile Trp Ser Cys Gly
        195                 200                 205

Val Ile Leu Tyr Val Leu Leu Ala Gly Phe Leu Pro Phe Gln Asp Asp
    210                 215                 220

Asn Leu Val Ala Leu Tyr Lys Lys Ile Tyr Arg Gly Asp Phe Lys Cys
225                 230                 235                 240

Pro Pro Trp Phe Ser Ser Glu Ala Arg Arg Leu Ile Thr Lys Leu Leu
                245                 250                 255

Asp Pro Asn Pro Asn Thr Arg Ile Thr Ile Ser Lys Ile Met Asp Ser
            260                 265                 270

Ser Trp Phe Lys Lys Pro Val Pro Lys Asn Leu Met Gly Lys Lys Arg
        275                 280                 285

Glu Glu Leu Asp Leu Glu Glu Lys Ile Lys Gln His Glu Gln Glu Val
    290                 295                 300
```

```
Ser Thr Thr Met Asn Ala Phe His Ile Ile Ser Leu Ser Glu Gly Phe
305                 310                 315                 320

Asp Leu Ser Pro Leu Phe Glu Glu Lys Lys Arg Glu Glu Lys Glu Leu
                325                 330                 335

Arg Phe Ala Thr Thr Arg Pro Ala Ser Ser Val Ile Ser Arg Leu Glu
            340                 345                 350

Asp Leu Ala Lys Ala Val Lys Phe Asp Val Lys Lys Ser Glu Thr Lys
        355                 360                 365

Val Arg Leu Gln Gly Gln Glu Lys Gly Arg Lys Gly Lys Leu Ala Ile
    370                 375                 380

Ala Ala Asp Leu Tyr Ala Val Thr Pro Ser Phe Leu Val Val Glu Val
385                 390                 395                 400

Lys Lys Asp Asn Gly Asp Thr Leu Glu Tyr Asn Gln Phe Cys Ser Lys
                405                 410                 415

Glu Leu Arg Pro Ala Leu Lys Asp Ile Val Trp Arg Thr Ser Pro Ala
            420                 425                 430

Glu Asn Pro Thr Leu Ala
        435
```

<210> 15
<211> 2543
<212> DNA
<213> Glycine max

<400> 15

```
caatgctctc ttcttccatt ttccttctca atccaccact ctaaccctcg cacgcttcgt   60
tcaattacaa aaatggacag atcaactggc cgtggtggtg gtggaagtgt ggacatgttt  120
ctccgaaatt ataagttggg aaaaacactc ggcattgggt cctttggcaa ggtgaaaatt  180
gctgagcatg tacggactgg tcataaagtt gctataaaga tccttaaccg ccacaagatt  240
aaaaacatgg aaatggaaga aaaagttaga agagaaatca aaattttaag attgtttatg  300
catcatcaca ttataagact atatgaggtt gtagaaaccc caacagacat atatgttgtt  360
atggagtatg tgaaatctgg agagctcttt gattacatag tagagaaggg tcggctgcaa  420
gaggatgaag cccgtcattt ttttcagcag ataatttctg gtgtggagta ctgtcacagg  480
aatatggtgg ttcatagaga cctgaagcct gagaatttac tcttggactc aaaatttaac  540
atcaagattg ctgattttgg gttgagcaac atcatgcgtg atggtcactt tcttaagaca  600
agttgtggaa gccctaatta tgcggctcca gaggttatct ctggaaaatt gtatgctgga  660
ccagaagtag atgtctggag ctgtggtgta attttatatg ctcttctctg tggcactctt  720
ccttttgatg atgaaaacat tcccaatctc ttcaaaaaaa taaagggtgg gatatacact  780
cttcctagtc atctatcacc tggtgctaga gatttgatac caaggatgct tgtggtggat  840
cccatgaaga ggatgaccat acctgagata cgccaacacc catggttcca agttcatcta  900
ccgcgttatt tagcagtgcc accaccagat acactgcaac aagccaaaaa gattgatgag  960
gagattcttc aggaagtggt taatatggga tttgacagga atcaattggt tgaatctctt 1020
agcaacagga tacaaaatga gggtactgta acatactatt tgttattgga caaccggttt 1080
cgtgtttcta gtggttatct tggagctgaa tttcaagaga caatggattc tggttttaac 1140
cgtatgcatt ccggcgaagt tgcttctcca gttgttggac accacagcac agggtatatg 1200
gattatcaag gggtaggaat gcggcaacag ttccctgttg agagaaaatg ggcccttggg 1260
cttcagtctc gagcccaacc acgtgaaata atgactgagg tccttaaagc tctacaagaa 1320
ttaaatgttt gttggaagaa gattggacac tataacatga agtgcagatg ggttgctggc 1380
actgctggtc atcatgaagg aatgattaac aattctctgc atagtaatca ttactttgga 1440
aatgattccg gcattattga aaatgaagct gtttctaagt caaatgtggt caagtttgaa 1500
gtgcagcttt acaaaactcg tgaggagaaa tatctgcttg atcttcaaag ggtccagggc 1560
```

```
ccacagtttc ttttcttgga tctgtgcgct gctttccttt cacagctacg tgttctctag 1620
tgaagctcac aagtcgcaaa ggagaccttg aggtgtcttg tgtggcatgt gaataagctt 1680
acgttgtaca tatcaatgcc ctttcagtat gggtatatct atctgtttgt actttttagc 1740
ataccccccag tttttatggg ttatcatgat tggggaaacc taaatatgga atttgctttc 1800
cttaatcatg gtgtctgact aggtctgtct ttttaaaagt gaattgaatt ttcttgatcc 1860
cgtgaaatcc ctgaggattt tatgaaagca ttcgcctcca acagagacat ccatttatgg 1920
tattgaattt tgtgttgtaa ttgactactt atgtctgtca cttggaagcc gaggggttgt 1980
ttgcctaggc tgcatcgaac atgagcttgc agataagcat tttctccccg cttccctact 2040
taattactgc ctcctcttat ggatgcatgt tatttgacaa aaaaaaaaat ctgaatggaa 2100
aagaaaagcg tcaccgtgca gatccgtgat tgagcaaaac gacgtcgtgc tgctacgcat 2160
cgcatcgaag ttgcaacaat ggcgagttcc gaatcgcaag agcatcatcc atggatctcg 2220
aatgcagttc ctttgttggt ggttcttcta attgcacttc acgtcttcgc tttggtgtat 2280
tggatttata gattggccac tgacaataag ccacaacagc agcagcagct acaacaacag 2340
caacagcaac atcaacagag aacaaaggct cactgaccga caacaacaac aacaacaaca 2400
acaaacgttc tcattcaatt tcatttttctt caaacaattg ttgtatgaaa ttgttaattg 2460
tgtgcagtaa aggatatgat ttttttgttt ttttggtata acagtgatga atgaagtttt 2520
gtttaatttt taaaaaaaaa aaa                                        2543
```

<210> 16
<211> 515
<212> PRT
<213> Glycine max

<400> 16

```
Met Asp Arg Ser Thr Gly Arg Gly Gly Gly Gly Ser Val Asp Met Phe
  1               5                  10                  15

Leu Arg Asn Tyr Lys Leu Gly Lys Thr Leu Gly Ile Gly Ser Phe Gly
             20                  25                  30

Lys Val Lys Ile Ala Glu His Val Arg Thr Gly His Lys Val Ala Ile
         35                  40                  45

Lys Ile Leu Asn Arg His Lys Ile Lys Asn Met Glu Met Glu Glu Lys
     50                  55                  60

Val Arg Arg Glu Ile Lys Ile Leu Arg Leu Phe Met His His His Ile
 65                  70                  75                  80

Ile Arg Leu Tyr Glu Val Val Glu Thr Pro Thr Asp Ile Tyr Val Val
                 85                  90                  95

Met Glu Tyr Val Lys Ser Gly Glu Leu Phe Asp Tyr Ile Val Glu Lys
            100                 105                 110

Gly Arg Leu Gln Glu Asp Glu Ala Arg His Phe Phe Gln Gln Ile Ile
        115                 120                 125

Ser Gly Val Glu Tyr Cys His Arg Asn Met Val Val His Arg Asp Leu
    130                 135                 140

Lys Pro Glu Asn Leu Leu Leu Asp Ser Lys Phe Asn Ile Lys Ile Ala
145                 150                 155                 160

Asp Phe Gly Leu Ser Asn Ile Met Arg Asp Gly His Phe Leu Lys Thr
                165                 170                 175

Ser Cys Gly Ser Pro Asn Tyr Ala Ala Pro Glu Val Ile Ser Gly Lys
            180                 185                 190
```

```
Leu Tyr Ala Gly Pro Glu Val Asp Val Trp Ser Cys Gly Val Ile Leu
        195                 200                 205

Tyr Ala Leu Leu Cys Gly Thr Leu Pro Phe Asp Asp Glu Asn Ile Pro
    210                 215                 220

Asn Leu Phe Lys Lys Ile Lys Gly Gly Ile Tyr Thr Leu Pro Ser His
225                 230                 235                 240

Leu Ser Pro Gly Ala Arg Asp Leu Ile Pro Arg Met Leu Val Val Asp
                245                 250                 255

Pro Met Lys Arg Met Thr Ile Pro Glu Ile Arg Gln His Pro Trp Phe
            260                 265                 270

Gln Val His Leu Pro Arg Tyr Leu Ala Val Pro Pro Pro Asp Thr Leu
        275                 280                 285

Gln Gln Ala Lys Lys Ile Asp Glu Glu Ile Leu Gln Glu Val Val Asn
    290                 295                 300

Met Gly Phe Asp Arg Asn Gln Leu Val Glu Ser Leu Ser Asn Arg Ile
305                 310                 315                 320

Gln Asn Glu Gly Thr Val Thr Tyr Tyr Leu Leu Leu Asp Asn Arg Phe
                325                 330                 335

Arg Val Ser Ser Gly Tyr Leu Gly Ala Glu Phe Gln Glu Thr Met Asp
            340                 345                 350

Ser Gly Phe Asn Arg Met His Ser Gly Glu Val Ala Ser Pro Val Val
        355                 360                 365

Gly His His Ser Thr Gly Tyr Met Asp Tyr Gln Gly Val Gly Met Arg
    370                 375                 380

Gln Gln Phe Pro Val Glu Arg Lys Trp Ala Leu Gly Leu Gln Ser Arg
385                 390                 395                 400

Ala Gln Pro Arg Glu Ile Met Thr Glu Val Leu Lys Ala Leu Gln Glu
                405                 410                 415

Leu Asn Val Cys Trp Lys Lys Ile Gly His Tyr Asn Met Lys Cys Arg
            420                 425                 430

Trp Val Ala Gly Thr Ala Gly His His Glu Gly Met Ile Asn Asn Ser
        435                 440                 445

Leu His Ser Asn His Tyr Phe Gly Asn Asp Ser Gly Ile Ile Glu Asn
    450                 455                 460

Glu Ala Val Ser Lys Ser Asn Val Val Lys Phe Glu Val Gln Leu Tyr
465                 470                 475                 480

Lys Thr Arg Glu Glu Lys Tyr Leu Leu Asp Leu Gln Arg Val Gln Gly
                485                 490                 495

Pro Gln Phe Leu Phe Leu Asp Leu Cys Ala Ala Phe Leu Ser Gln Leu
            500                 505                 510
```

Arg Val Leu
515

<210> 17
<211> 1869
<212> DNA
<213> Glycine max

<400> 17
gcacgaggtc tggttgcata gcattggttg gtagttgtct caaaaatctc ttcttgccct 60
ttggccataa tcaaaagcca agacactgtt catacagctg ctcaattatc aagccaacct 120
tgctcggttc cactgcagaa tttcagttta ttcttatcta gctcaattct ggttgtgggt 180
ttatctctta ctggaagaca gactttgagg tagactcctt ataagtgcgc agaagttcaa 240
gtgtagagaa tgagtcagcc taagattaaa cgccgagttg gtaaatacga ggtggggagg 300
accattggtg aaggtacatt tgcaaaggtg aaatttgcaa ggaactctga gacaggagag 360
cccgtggctc ttaaaattct tgacaaggag aaggtgctaa agcacaagat ggctgagcag 420
atcaggagag aagtagctac aatgaaacta atcaagcatc caaatgttgt tcgattgtat 480
gaggtcatgg gaagcaagac caaaatatat attgttttgg agtttgtaac tgggggggaa 540
ctctttgaca aaattgtaaa ccatggaagg atgagtgaaa atgaagcacg tagatatttc 600
cagcagctta taaatgctgt tgattattgc catagcaggg gtgtctacca cagagacctg 660
aagccagaaa atttgctatt agatacttat gggaacctta aagtttctga ttttggtttg 720
agtgccctct cccagcaagt tagggatgat ggacttcttc atactacatg tggcactcca 780
aattatgttg ctcctgaggt ccttaacgat agaggctatg atggggcaac tgcagacttg 840
tggtcatgtg gggttattct ctttgtattg gttgcaggtt acttgccttt cgacgaccct 900
aatcttatga acctgtataa aaagatctca gctgctgaat ttacttgccc cccatggctt 960
tctttcactg ccaggaaatt gattacacga atcttggatc cagatcccac cactcgtatc 1020
actatacctg agattttgga tgatgaatgg tttaagaaag aatataagcc tcccatttt 1080
gaggagaatg gggaaatcaa cctcgatgat gttgaagctg tctttaaaga ctctgaagag 1140
caccatgtga cagagaaaaa agaagagcag cctacagcca tgaatgcatt tgagttaatc 1200
tccatgtcca aaggactgaa ccttgaaaac ttgtttgata ctgagcaggg atttaaaagg 1260
gaaacaagat tcacctcaaa atcccctgcg gatgagataa tcaacaagat tgaggaagcc 1320
gcaaaacctc ttggctttga tgtgcagaag aaaaattaca agatgaggct tgcaaatgtg 1380
aaagctggaa ggaagggaaa ccttaatgtt gccacagaga tatttcaagt ggcaccttct 1440
cttcacatgg tagaggtacg gaaggcaaaa ggagatacat tggagttcca taagttctac 1500
aagaaacttt caacaagcct ggatgatgtt gtttggaaaa cagaagatga tatgcaaatg 1560
cgagaaacaa agtgatgtgg atattattat cattgtctat taagtgtaat tttcttcgtg 1620
tctgaggttt tactattttc caatttcttc attcgttata ttcctccccc gtaggtttgt 1680
ttggacatta attacatagt actcatttat tgcataccat gctattattt tttgaaagca 1740
tgcagagttc atgtaagaat tttactcatc caacagtcgc ggttatgttc atgaaacaaa 1800
aaattgtaag aaatttgtat attgtatata tctatctatt tatatctttt caaaaaaaaa 1860
aaaaaaaaa 1869

<210> 18
<211> 441
<212> PRT
<213> Glycine max

<400> 18
Met Ser Gln Pro Lys Ile Lys Arg Arg Val Gly Lys Tyr Glu Val Gly
1 5 10 15

Arg Thr Ile Gly Glu Gly Thr Phe Ala Lys Val Lys Phe Ala Arg Asn
20 25 30

Ser Glu Thr Gly Glu Pro Val Ala Leu Lys Ile Leu Asp Lys Glu Lys
35 40 45

Val Leu Lys His Lys Met Ala Glu Gln Ile Arg Arg Glu Val Ala Thr
50 55 60

```
Met Lys Leu Ile Lys His Pro Asn Val Val Arg Leu Tyr Glu Val Met
 65                  70                  75                  80

Gly Ser Lys Thr Lys Ile Tyr Ile Val Leu Glu Phe Val Thr Gly Gly
                 85                  90                  95

Glu Leu Phe Asp Lys Ile Val Asn His Gly Arg Met Ser Glu Asn Glu
            100                 105                 110

Ala Arg Arg Tyr Phe Gln Gln Leu Ile Asn Ala Val Asp Tyr Cys His
        115                 120                 125

Ser Arg Gly Val Tyr His Arg Asp Leu Lys Pro Glu Asn Leu Leu Leu
    130                 135                 140

Asp Thr Tyr Gly Asn Leu Lys Val Ser Asp Phe Gly Leu Ser Ala Leu
145                 150                 155                 160

Ser Gln Gln Val Arg Asp Asp Gly Leu Leu His Thr Thr Cys Gly Thr
                165                 170                 175

Pro Asn Tyr Val Ala Pro Glu Val Leu Asn Asp Arg Gly Tyr Asp Gly
            180                 185                 190

Ala Thr Ala Asp Leu Trp Ser Cys Gly Val Ile Leu Phe Val Leu Val
        195                 200                 205

Ala Gly Tyr Leu Pro Phe Asp Asp Pro Asn Leu Met Asn Leu Tyr Lys
    210                 215                 220

Lys Ile Ser Ala Ala Glu Phe Thr Cys Pro Pro Trp Leu Ser Phe Thr
225                 230                 235                 240

Ala Arg Lys Leu Ile Thr Arg Ile Leu Asp Pro Asp Pro Thr Thr Arg
                245                 250                 255

Ile Thr Ile Pro Glu Ile Leu Asp Asp Glu Trp Phe Lys Lys Glu Tyr
            260                 265                 270

Lys Pro Pro Ile Phe Glu Glu Asn Gly Glu Ile Asn Leu Asp Asp Val
        275                 280                 285

Glu Ala Val Phe Lys Asp Ser Glu Glu His His Val Thr Glu Lys Lys
    290                 295                 300

Glu Glu Gln Pro Thr Ala Met Asn Ala Phe Glu Leu Ile Ser Met Ser
305                 310                 315                 320

Lys Gly Leu Asn Leu Glu Asn Leu Phe Asp Thr Glu Gln Gly Phe Lys
                325                 330                 335

Arg Glu Thr Arg Phe Thr Ser Lys Ser Pro Ala Asp Glu Ile Ile Asn
            340                 345                 350

Lys Ile Glu Glu Ala Ala Lys Pro Leu Gly Phe Asp Val Gln Lys Lys
        355                 360                 365

Asn Tyr Lys Met Arg Leu Ala Asn Val Lys Ala Gly Arg Lys Gly Asn
    370                 375                 380
```

Leu Asn Val Ala Thr Glu Ile Phe Gln Val Ala Pro Ser Leu His Met
385 390 395 400

Val Glu Val Arg Lys Ala Lys Gly Asp Thr Leu Glu Phe His Lys Phe
405 410 415

Tyr Lys Lys Leu Ser Thr Ser Leu Asp Asp Val Val Trp Lys Thr Glu
420 425 430

Asp Asp Met Gln Met Arg Glu Thr Lys
435 440

<210> 19
<211> 817
<212> DNA
<213> Triticum aestivum

<400> 19
gcgctctcgt cgcctcccca cccattccag cggttgaccg cgggattcgg ccagtgaaaa 60
tggaagggaa cactagagga ggtgggcatt ctgacgcatt aaagaactac aatgtgggca 120
gaacattagg tataggcaca tttggaaaag tgaggattgc agagcataag catacagggc 180
ataaagttgc tataaagatt ctgaaccgtc gtcaaatgag aactatggaa atggaggaga 240
aagcaaagag agagatcaag atattgaggt tgttcatcca ccctcatatc atccggcttt 300
atgaggtcat ttacacacct acagatatat ttgttgtgat ggaatattgc aagtatggtg 360
agctattcga ctgcattgtt gagaaagggc ggttacagga agatgaggct cgtcgaatct 420
tccagcagat tatatctggt gttgaatact gccacagaaa catggttgct catcgtgatc 480
taaagccaga gaacctgtta cttgattcca aatacaatgt gaaacttgcc gactttgggt 540
taagtaatgt catgcatgat ggccattttc tgaagactag ctgcgggagt ccaaactatg 600
ctgcaccaga ggttatctca ggtaaattat acgctggacc tgaggttgat gtttggagct 660
gcggggtgat actttatgct cttctttgtg gcactcttcc atttgatgat gacaatattc 720
ccaaactgtt caaaaagata aagggaggca tctatatcct tccaagtcat ttatctgctc 780
ctgcaaggga ttgatccaag aatgcttgtt gttgatc 817

<210> 20
<211> 244
<212> PRT
<213> Triticum aestivum

<400> 20
Met Glu Gly Asn Thr Arg Gly Gly Gly His Ser Asp Ala Leu Lys Asn
1 5 10 15

Tyr Asn Val Gly Arg Thr Leu Gly Ile Gly Thr Phe Gly Lys Val Arg
20 25 30

Ile Ala Glu His Lys His Thr Gly His Lys Val Ala Ile Lys Ile Leu
35 40 45

Asn Arg Arg Gln Met Arg Thr Met Glu Met Glu Glu Lys Ala Lys Arg
50 55 60

Glu Ile Lys Ile Leu Arg Leu Phe Ile His Pro His Ile Ile Arg Leu
65 70 75 80

Tyr Glu Val Ile Tyr Thr Pro Thr Asp Ile Phe Val Val Met Glu Tyr
85 90 95

Cys Lys Tyr Gly Glu Leu Phe Asp Cys Ile Val Glu Lys Gly Arg Leu
100 105 110

Gln Glu Asp Glu Ala Arg Arg Ile Phe Gln Gln Ile Ile Ser Gly Val
115 120 125

Glu Tyr Cys His Arg Asn Met Val Ala His Arg Asp Leu Lys Pro Glu
130 135 140

Asn Leu Leu Leu Asp Ser Lys Tyr Asn Val Lys Leu Ala Asp Phe Gly
145 150 155 160

Leu Ser Asn Val Met His Asp Gly His Phe Leu Lys Thr Ser Cys Gly
165 170 175

Ser Pro Asn Tyr Ala Ala Pro Glu Val Ile Ser Gly Lys Leu Tyr Ala
180 185 190

Gly Pro Glu Val Asp Val Trp Ser Cys Gly Val Ile Leu Tyr Ala Leu
195 200 205

Leu Cys Gly Thr Leu Pro Phe Asp Asp Asp Asn Ile Pro Lys Leu Phe
210 215 220

Lys Lys Ile Lys Gly Gly Ile Tyr Ile Leu Pro Ser His Leu Ser Ala
225 230 235 240

Pro Ala Arg Asp

<210> 21
<211> 2006
<212> DNA
<213> Triticum aestivum

<400> 21
ctccgcgccg ccgctgccgc tacgcctctc cccgggaagc ctcgccggcg gccaggtgga 60
agatggagac aggcggcaaa gatggcaacc ctttgaagaa ttaccgtatt gggaagaccc 120
tggggattgg ttcgttcggg aaggtcaaga ttgccgagca tataaaaact ggtcacaagg 180
tggccgtcaa gatccttaac cgccggaaaa tcaaaaacat ggagatggaa gagaaagtga 240
aaagagagat caagatatta agattattca tgcacccaca tatcatccgc ctttatgaag 300
tgatagaggc accagctgat atttatgtgg ttatggagta tgttaagtct ggtgaattgt 360
ttgattacat tgttgagaaa ggtaggctac aggaggaaga ggcccgccgt ttctttcaac 420
agatcatatc tggtgttcaa tattgccaca ggaacatggt ggtgcaccgc gatctaaagc 480
cggagaacct tcttttggac aataattgtg atgttaagat tgcggatttt ggcttaagta 540
atgttatgcg tgacggccac tttcttaaga caagttgtgg tagcccaaat tatgcagctc 600
cggaggttat atctggaaaa ctgtacgctg ggcctgaagt tgatgtatgg agctgcggtg 660
ttattcttta tgctcttcta tgtggtactc ttccatttga tgatgagaac atacccaacc 720
tttttaagaa aataaagggt ggaatatata cccttccaag ccatttatca ggcccagcaa 780
gggatttgat tccaaggatg ctagttgttg atcctatgaa gaggataacc attcgtgaaa 840
tacgcgagca tccatggttt gaagctcaac tcccacgata tttagccgtg cctccaccag 900
atactgcaca acaagttaaa aagattgatg aagaatctct tgttaaagtt atcagtctgg 960
gatttgacaa aaacctgctg gttgaatcaa ttcataatag attgcaaaat gaggcaacag 1020
ttgcatatta tttgtttttg gataataaga gtcgcacaac aactggctat cttggagctg 1080
ggtatcaaga agctatggaa tcgtctttct cacccattac tccaagtgaa acacaaagtc 1140
cagctcatgg aaatcggcaa caaccatata tggaatctcc agttggcttg agaccacatt 1200
ttccagctga taggaaatgg gctcttgggc ttcagtctcg agcacatcca agagaagtta 1260
tgactgaagt gctgaaggct ctgcaagaac tgaatgtata ctggaaaaaa attggacact 1320
ataacatgaa atgtagatgg agtcctcctg gctttcccgg tcaggagaat atgaatcata 1380
ccaattataa cttcagtgca gagcctattg aaaccgacga cctgggtgac aagttaaatt 1440
taattaagtt cgaacttcag ctttacaaaa caagagatga gaaatacctt ctggatttgc 1500
aaagggcgag cgggccgcat ctcctctttc ttgatctatg tgccgccttt ctagctcagc 1560
tgagagtctt ttgataccag atgtgcccga ggaatgtatg ttgtatcact ctaaagagat 1620

gtaaatagca agctttctcc agcggatcaa agtcgtggag tatgtagaca tgcggagctg 1680
ttgtgtgctt atttcggcgc ctatatgctg aatttagacc tggcaggggc gggcaagtga 1740
agcaagcaag gaactattgc catcaggtta tttccagctg ccgccaaagg cactaggata 1800
tagaagtatt actgattaat cctatattgg ccccttggga catactccta ctctactgct 1860
gtttacttgc atgtaatttt tactgtctgg gtctccagac cagaccacgt acacgaataa 1920
tttcttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1980
aaaaaaaaaa aaaaaaaaaa aaaaaa 2006

<210> 22
<211> 523
<212> PRT
<213> Triticum aestivum

<400> 22

Pro	Arg	Arg	Arg	Cys	Arg	Tyr	Ala	Ser	Pro	Arg	Glu	Ala	Ser	Pro	Ala
1				5					10					15	
Ala	Arg	Trp	Lys	Met	Glu	Thr	Gly	Gly	Lys	Asp	Gly	Asn	Pro	Leu	Lys
			20					25					30		
Asn	Tyr	Arg	Ile	Gly	Lys	Thr	Leu	Gly	Ile	Gly	Ser	Phe	Gly	Lys	Val
		35					40					45			
Lys	Ile	Ala	Glu	His	Ile	Lys	Thr	Gly	His	Lys	Val	Ala	Val	Lys	Ile
	50					55					60				
Leu	Asn	Arg	Arg	Lys	Ile	Lys	Asn	Met	Glu	Met	Glu	Glu	Lys	Val	Lys
65					70					75					80
Arg	Glu	Ile	Lys	Ile	Leu	Arg	Leu	Phe	Met	His	Pro	His	Ile	Ile	Arg
				85					90					95	
Leu	Tyr	Glu	Val	Ile	Glu	Ala	Pro	Ala	Asp	Ile	Tyr	Val	Val	Met	Glu
			100					105					110		
Tyr	Val	Lys	Ser	Gly	Glu	Leu	Phe	Asp	Tyr	Ile	Val	Glu	Lys	Gly	Arg
		115					120					125			
Leu	Gln	Glu	Glu	Glu	Ala	Arg	Arg	Phe	Phe	Gln	Gln	Ile	Ile	Ser	Gly
	130					135					140				
Val	Gln	Tyr	Cys	His	Arg	Asn	Met	Val	Val	His	Arg	Asp	Leu	Lys	Pro
145					150					155					160
Glu	Asn	Leu	Leu	Leu	Asp	Asn	Asn	Cys	Asp	Val	Lys	Ile	Ala	Asp	Phe
				165					170					175	
Gly	Leu	Ser	Asn	Val	Met	Arg	Asp	Gly	His	Phe	Leu	Lys	Thr	Ser	Cys
			180					185					190		
Gly	Ser	Pro	Asn	Tyr	Ala	Ala	Pro	Glu	Val	Ile	Ser	Gly	Lys	Leu	Tyr
		195					200					205			
Ala	Gly	Pro	Glu	Val	Asp	Val	Trp	Ser	Cys	Gly	Val	Ile	Leu	Tyr	Ala
	210					215					220				
Leu	Leu	Cys	Gly	Thr	Leu	Pro	Phe	Asp	Asp	Glu	Asn	Ile	Pro	Asn	Leu
225					230					235					240

```
Phe Lys Lys Ile Lys Gly Gly Ile Tyr Thr Leu Pro Ser His Leu Ser
                245                 250                 255

Gly Pro Ala Arg Asp Leu Ile Pro Arg Met Leu Val Val Asp Pro Met
            260                 265                 270

Lys Arg Ile Thr Ile Arg Glu Ile Arg Glu His Pro Trp Phe Glu Ala
        275                 280                 285

Gln Leu Pro Arg Tyr Leu Ala Val Pro Pro Pro Asp Thr Ala Gln Gln
    290                 295                 300

Val Lys Lys Ile Asp Glu Glu Ser Leu Val Lys Val Ile Ser Leu Gly
305                 310                 315                 320

Phe Asp Lys Asn Leu Leu Val Glu Ser Ile His Asn Arg Leu Gln Asn
                325                 330                 335

Glu Ala Thr Val Ala Tyr Tyr Leu Phe Leu Asp Asn Lys Ser Arg Thr
            340                 345                 350

Thr Thr Gly Tyr Leu Gly Ala Gly Tyr Gln Glu Ala Met Glu Ser Ser
        355                 360                 365

Phe Ser Pro Ile Thr Pro Ser Glu Thr Gln Ser Pro Ala His Gly Asn
    370                 375                 380

Arg Gln Gln Pro Tyr Met Glu Ser Pro Val Gly Leu Arg Pro His Phe
385                 390                 395                 400

Pro Ala Asp Arg Lys Trp Ala Leu Gly Leu Gln Ser Arg Ala His Pro
                405                 410                 415

Arg Glu Val Met Thr Glu Val Leu Lys Ala Leu Gln Glu Leu Asn Val
            420                 425                 430

Tyr Trp Lys Lys Ile Gly His Tyr Asn Met Lys Cys Arg Trp Ser Pro
        435                 440                 445

Pro Gly Phe Pro Gly Gln Glu Asn Met Asn His Thr Asn Tyr Asn Phe
    450                 455                 460

Ser Ala Glu Pro Ile Glu Thr Asp Asp Leu Gly Asp Lys Leu Asn Leu
465                 470                 475                 480

Ile Lys Phe Glu Leu Gln Leu Tyr Lys Thr Arg Asp Glu Lys Tyr Leu
                485                 490                 495

Leu Asp Leu Gln Arg Ala Ser Gly Pro His Leu Leu Phe Leu Asp Leu
            500                 505                 510

Cys Ala Ala Phe Leu Ala Gln Leu Arg Val Phe
        515                 520
```

<210> 23
<211> 512
<212> DNA
<213> Zea mays

<400> 23
gagcagctcc cctgcccctc gcagcggcta ctctacaggt ctagcgactc tttcgccatc 60
catagaggga ggaggcgcgg cggagatggt gggcggtggc ggcggcgggc cgctgcggcg 120
ggtgggcaag tacgaggtgg gacgcaccat cggggaaggc accttcgcca aggtcaagtt 180
cgcgcagaac accgagaccg gggagagcgt cgccatgaag gtgctcgacc gctcctccat 240
cctcaagaac aagatggccg aacagattaa gagagaaata tccataatga agcttgtcag 300
gcatcccaat gtcgttaggc tacacgaggt tttggcaagc cggaagaaga tatttataat 360
tctggagttc atcactggcg gcgagctatt cgataaaatt attcgtcatg ggagactcag 420
tgaagcagat gcccgcagat actttcagca gcttattgat ggtgttgatt tttgtcacaa 480
gaaaggagtc taccatcgag acttaaagcc tg 512

<210> 24
<211> 132
<212> PRT
<213> Zea mays

<400> 24
Arg Arg Val Gly Lys Tyr Glu Val Gly Arg Thr Ile Gly Glu Gly Thr
1 5 10 15

Phe Ala Lys Val Lys Phe Ala Gln Asn Thr Glu Thr Gly Glu Ser Val
20 25 30

Ala Met Lys Val Leu Asp Arg Ser Ser Ile Leu Lys Asn Lys Met Ala
35 40 45

Glu Gln Ile Lys Arg Glu Ile Ser Ile Met Lys Leu Val Arg His Pro
50 55 60

Asn Val Val Arg Leu His Glu Val Leu Ala Ser Arg Lys Lys Ile Phe
65 70 75 80

Ile Ile Leu Glu Phe Ile Thr Gly Gly Glu Leu Phe Asp Lys Ile Ile
85 90 95

Arg His Gly Arg Leu Ser Glu Ala Asp Ala Arg Arg Tyr Phe Gln Gln
100 105 110

Leu Ile Asp Gly Val Asp Phe Cys His Lys Lys Gly Val Tyr His Arg
115 120 125

Asp Leu Lys Pro
130

<210> 25
<211> 552
<212> DNA
<213> Glycine max

<220>
<221> unsure
<222> (385)

<400> 25
gtctggttgc atagcattgg ttggtagttg tctcaaaaat ctcttcttgc cctttggcca 60
taatcaaaag ccaagacact gttcatacag ctgctcaatt atcaagccaa ccttgctcgg 120
ttccactgca gaatttcagt ttattcttat ctagctcaat tctggttgtg ggtttatctc 180
ttactggaag acagactttg aggtagactc cttataagtg cgcagaagtt caagtgtaga 240
gaatgagtca gcctaagatt aaacgccgag ttggtaaata cgaggtgggg aggaccattg 300
gtgaaggtac atttgcaaag gtgaaatttg caaggaactc tgagacagga gagccgtggc 360

tcttaaaatt cttgacaagg agaangtgct aaagcacaag atggctgagc agatcaggag 420
agaagtagct acaatgaaac taatcaagca tccaaatgtt gttcgattgt atgaagtcat 480
gggaagcaag acaaatatat aatgttttgg agttgtactg ggggggaacc cttgcaaatt 540
gtaaccatgg aa 552

<210> 26
<211> 77
<212> PRT
<213> Glycine max

<220>
<221> UNSURE
<222> (39)

<400> 26
Val Gly Lys Tyr Glu Val Gly Arg Thr Ile Gly Glu Gly Thr Phe Ala
1 5 10 15

Lys Val Lys Phe Ala Arg Asn Ser Glu Thr Gly Glu Pro Trp Leu Leu
20 25 30

Lys Phe Leu Thr Arg Arg Xaa Val Leu Lys His Lys Met Ala Glu Gln
35 40 45

Ile Arg Arg Glu Val Ala Thr Met Lys Leu Ile Lys His Pro Asn Val
50 55 60

Val Arg Leu Tyr Glu Val Met Gly Ser Lys Thr Asn Ile
65 70 75

<210> 27
<211> 391
<212> DNA
<213> Triticum aestivum

<220>
<221> unsure
<222> (179)

<220>
<221> unsure
<222> (236)

<220>
<221> unsure
<222> (240)

<220>
<221> unsure
<222> (297)

<220>
<221> unsure
<222> (316)

<220>
<221> unsure
<222> (331)

<220>
<221> unsure
<222> (336)

<220>
<221> unsure
<222> (338)

<220>
<221> unsure
<222> (344)

<220>
<221> unsure
<222> (357)

<220>
<221> unsure
<222> (361)

<220>
<221> unsure
<222> (371)

<220>
<221> unsure
<222> (381)

<220>
<221> unsure
<222> (386)

<400> 27
ctccgcgccg ccgctgccgc tacgcctctc cccgggaagc ctcgccggcg gccaggtgga 60
agatggagac aggcggcaaa gatggcaacc ctttgaagaa ttaccgtatt gggaagaccç 120
tggggattgg ttcgttcggg aaggtcaaga ttgccgagca tataaaaact ggtcacaang 180
tggccgtcaa gatccttaac cgccggcaaa tcaaaaacat ggcgatggaa gagaangtgn 240
caagagagat caagatatta agattattca tgcacccaca tatcatccgc ctttatnaag 300
tgatagaggc accagntgat atttatgtgg ntatgnanta tgtnaaagtc cggtganttg 360
nttgattata ntgtttctaa ngtctntata t 391

<210> 28
<211> 85
<212> PRT
<213> Triticum aestivum

<220>
<221> UNSURE
<222> (29)

<220>
<221> UNSURE
<222> (48)

<220>
<221> UNSURE
<222> (50)

<220>
<221> UNSURE
<222> (69)

<220>
<221> UNSURE
<222> (75)

<220>
<221> UNSURE
<222> (80)

<220>
<221> UNSURE
<222> (82)

<400> 28

Leu Lys Asn Tyr Arg Ile Gly Lys Thr Leu Gly Ile Gly Ser Phe Gly
1 5 10 15

Lys Val Lys Ile Ala Glu His Ile Lys Thr Gly His Xaa Val Ala Val
20 25 30

Lys Ile Leu Asn Arg Arg Gln Ile Lys Asn Met Ala Met Glu Glu Xaa
35 40 45

Val Xaa Arg Glu Ile Lys Ile Leu Arg Leu Phe Met His Pro His Ile
50 55 60

Ile Arg Leu Tyr Xaa Val Ile Glu Ala Pro Xaa Asp Ile Tyr Val Xaa
65 70 75 80

Met Xaa Tyr Val Lys
85